On Variable Bandwidth Kernel Density And Regression Estimation

Janet Nakarmi
University of Mississippi

ON VARIABLE BANDWIDTH KERNEL DENSITY AND REGRESSION ESTIMATION

DISSERTATION

A Dissertation
presented in partial fulfillment of requirements
for the degree of Doctor of Philosophy
in the Department of Mathematics
The University of Mississippi

by

JANET NAKARMI

August 2016

ABSTRACT

We study the ideal variable bandwidth kernel density estimator introduced by McKay [16, 17] and the plug-in practical version of the variable bandwidth kernel density estimator with two sequences of bandwidths as in [10]. We estimate the variance of the variable bandwidth kernel density estimator. Based on the exact formula of the bias and the variance of the variable bandwidth kernel density estimator, we develop the optimal bandwidth selection of the true variable bandwidth kernel density estimator. Furthermore, we present the central limit theorem of the true variable bandwidth kernel density estimator. We also propose a new variable bandwidth kernel regression estimator and estimate the bias and propose the central limit theorems for its ideal and true versions.

For the one dimensional case, the order of the bias and variance is same for the variable bandwidth kernel density estimator and for the proposed variable bandwidth kernel regression estimator. Since we use the order of the bias and variance to find the optimal bandwidth, the optimal bandwidth for these estimators are also the same. Comparing the integrated mean square error of the variable bandwidth kernel density estimator (the variable bandwidth kernel regression estimator) with the classical kernel density estimator (the Nadaraya-Watson estimator), we find that the variable bandwidth kernel estimators have a faster rate of convergence. Furthermore, we prove that these variable bandwidth kernel estimators converge to normal distribution.

DEDICATION

This dissertation is dedicated to my parents Narayan Lal and Purna Laxmi Nakarmi. Without their support, love and encouragement, I would not have been able to come this far. I would also like to dedicate this work to my best friend, Himani Vaidya for being there for me throughout the entire graduate program.

ACKNOWLEDGEMENTS

I would like to thank my advisor Dr. Hailin Sang for his continuous support, patience and vast knowledge. His guidance throughout the whole process of writing this dissertation has helped me immensely. I could not have imagined having a better advisor for my Ph.D research.

Besides my advisor, I would also like to thank Dr. Xin Dang, Dr. Erwin Miña Diaz, Dr. Yixin Chen, and Dr. Martial Longla for their insightful comments and for giving their time to serve on my committee. I would like to thank Dr. James Reid for his support and vast knowledge in LaTeX. A special thanks goes to Dr. Kayla Harville, Colleen Gilman, Casandra Jenkins, Shelia Lewis, and Kay Phillips, who were always there to help and encourage me anytime that I needed it. I greatly appreciate the support of the entire Department of Mathematics and cherish the friendship and support that I have received from all of its members.

TABLE OF CONTENTS

LIST OF FIGURES

1 INTRODUCTION

1.1 CLASSICAL KERNEL DENSITY ESTIMATOR

Suppose that X_i, for $i = 1, \cdots, n, n \in \mathbb{N}$, are independent and identically distributed (i.i.d.) observations with density function $f(t)$, $t \in \mathbb{R}^d$. The goal of nonparametric density estimation is to estimate f with as few assumptions about f as possible. One of the well known estimators of f is the classical kernel density estimator, which we will denote by

$$\hat{f}(t; h_n) = \frac{1}{nh_n^d} \sum_{i=1}^{n} K\left(\frac{t - X_i}{h_n}\right), \tag{1.1.1}$$

where h_n is the bandwidth sequence satisfying $h_n \to 0, nh_n^d \to \infty$ and K is a kernel function. A kernel function, K, is any function that satisfies $\int K(x)dx = 1$. The kernel K satisfies the following conditions to get the order of the bias and variance of the classical kernel density estimator: $K(x) \geq 0$, $\int xK(x)dx = 0$, and $\sigma_K^2 = \int x^2K(x)dx > 0$. The variance of (1.1.1) has order $O((nh_n^d)^{-1})$ and the bias has order $O(h_n^2)$ when $f(t)$ has bounded second order partial derivatives. See [25] and [27] for the literature on kernel density estimation.

Throughout this dissertation we use the following definitions and notations.

Definition 1.1.2. *I)* $a_n = o(b_n)$ *iff* $a_n/b_n \to 0$.

II) $a_n = O(b_n)$ *iff there exists constants* S *and* n_0 *such that* $|a_n| \leq Sb_n$ *for* $n \geq n_0$.

III) $X_n = O_p(a_n)$ *iff for every* $\epsilon > 0$ *there exists constants* C_ϵ *and* n_ϵ *such that* $\mathbb{P}(|X_n| \leq C_\epsilon a_n) > 1 - \epsilon$ *for every* $n \geq n_\epsilon$.

IV) $X_n = o_p(a_n)$ *iff* $X_n/a_n \xrightarrow{p} 0$.

V) $X_n = O_{a.s}(a_n)$ *iff there exists a constant* B *such that* $|X_n| \leq Ba_n$ *a.s. for* n *large enough.*

VI) $X_n = o_{a.s}(a_n)$ *iff there exists a sequence* $\delta_n \to 0$ *such that* $|X_n| \leq \delta_n a_n$ *a.s.*

1.2 VARIABLE BANDWIDTH KERNEL DENSITY ESTIMATOR

We study the following multidimensional version of the variable bandwidth kernel density estimator proposed in [16] and [17]:

$$\bar{f}(t; h_n) = \frac{1}{nh_n^d} \sum_{i=1}^{n} \alpha^d(f(X_i)) K(h_n^{-1} \alpha(f(X_i))(t - X_i)), \tag{1.2.1}$$

with

$$\alpha(s) := cp^{1/2}(s/c^2), \tag{1.2.2}$$

where c is a fixed number with $0 < c < \infty$ and the function p has at least fourth order derivative and satisfies the following conditions:

$$p(x) \geq 1 \text{ for all } x \in \mathbb{R}, \tag{1.2.3}$$

$$p(x) = x \text{ for all } x \geq t_0 \text{ for some } 1 \leq t_0 < \infty. \tag{1.2.4}$$

The study of variable bandwidth kernel density estimation goes back to [1]. Abramson [1] proposed the estimator

$$f_A(t; h_n) = \frac{1}{nh_n^d} \sum_{i=1}^{n} \gamma^d(t, X_i) K(h_n^{-1} \gamma(t, X_i)(t - X_i)), \tag{1.2.5}$$

where $\gamma(t, s) = (f(s) \vee f(t)/10)^{1/2}$. The true bandwidth, $h_n/\gamma(t, X_i)$, at each observation X_i is inversely proportional to $f^{1/2}(X_i)$ if $f(X_i) \geq f(t)/10$ (which is the square root law). Notice that (1.2.1) also has the square root law since $\alpha(f(X_i)) = f^{1/2}(X_i)$ if $f(X_i) \geq t_0 c^2$ by the definition of the function $p(x)$. The estimator (1.2.1) or (1.2.5) has clipping procedure in (1.2.2) or $\gamma(t, s)$ since they make the true bandwidth $h_n/\alpha(f(X_i)) \geq h_n/c$ or

$h_n/\gamma(t, X_i) \geq 10^{1/2} h_n / f(t)^{1/2}$. The clipping procedures prevent too much contribution to the density estimation at t if the observation X_i is too far away from t. Abramson showed that this square root law and the clipping procedure improve the bias from the order of h_n^2 to the order of h_n^4 for the estimator (1.2.5), while at the same time keep the variance at the order of $(nh_n^d)^{-1}$ if $f(t) \neq 0$ and $f(t)$ has fourth order continuous derivatives at t. However, this variable bandwidth estimator (1.2.5) is not a density function since the integral of $f_A(t; h_n)$ over t is not 1.

Terrell & Scott [26] and McKay [17] showed that a modification of the Abramson estimator without the 'clipping filter' $(f(t)/10)^{1/2}$ on $f^{1/2}(X_i)$ studied in [13], namely,

$$f_{HM}(t; h_n) = \frac{1}{nh_n^d} \sum_{i=1}^{n} f^{d/2}(X_i) K(h_n^{-1} f^{1/2}(X_i)(t - X_i)), \tag{1.2.6}$$

which has integral 1 and thus is a true probability density, may have bias of order much larger than h_n^4. Therefore, the clipping is necessary for bias reduction. In the case $d = 1$, Hall *et al.* [12] proposed the estimator

$$f_{HHM}(t; h_n) = \frac{1}{nh_n} \sum_{i=1}^{n} K\left(\frac{t - X_i}{h_n} f^{1/2}(X_i)\right) f^{1/2}(X_i) I(|t - X_i| < h_n B), \tag{1.2.7}$$

where B is a fixed constant. This estimator is non-negative and achieves the desired bias reduction but, like Abramson's estimator, it does not integrate to 1. See also [21] for a similar estimator.

In conclusion, it seems that the estimator (1.2.1) has all the advantages: it is a true density function with square root law and smooth clipping procedure. However, notice that this estimator and all the other variable bandwidth kernel density estimators are not applicable in practice since they all include the studied density function f. They are called ideal estimators in the literature. Hall & Marron [13] studied the true density estimator

$$\hat{f}_{HM}(t; h_{1,n}, h_{2,n}) = \frac{1}{nh_{2,n}^d} \sum_{i=1}^{n} K\left(\frac{t - X_i}{h_{2,n}} \hat{f}^{1/2}(X_i; h_{1,n})\right) \hat{f}^{d/2}(X_i; h_{1,n}),$$

by plugging in a pilot estimator, the classical estimator (1.1.1), into the estimator (1.2.6). They used the Taylor expansion of $K\left(\frac{t-X_i}{h_{2,n}}\hat{f}^{1/2}(X_i;h_{1,n})\right)$ at $K\left(\frac{t-X_i}{h_{2,n}}f^{1/2}(X_i)\right)$ to prove that the difference between the true estimator $\hat{f}_{HM}(t;h_{1,n},h_{2,n})$ and the ideal version (1.2.6) has pointwise asymptotic convergence rate $O_P(n^{-4/(8+d)})$. By applying this Taylor decomposition, McKay [17] studied convergence in probability and pointwise of plug-in true estimator of (1.2.1). Giné & Sang [9, 10] studied plug-in true estimators of (1.2.7) and (1.2.1) for one and d-dimensional observations. They proved that the difference between the true estimator and the true value ($f(t)$) converges uniformly over a data adaptive region at the rate of $O_{a.s.}((\log n/n)^{4/(8+d)})$ by applying empirical process techniques. The true estimator in [10] has the following form:

$$\hat{f}(t;h_{1,n},h_{2,n}) = \frac{1}{nh_{2,n}^d}\sum_{i=1}^{n} K\left(\frac{t-X_i}{h_{2,n}}\alpha(\hat{f}(X_i;h_{1,n}))\right)\alpha^d(\hat{f}(X_i;h_{1,n})). \tag{1.2.8}$$

Jones *et al.* [15] studied ideal estimators with bias order h_n^6. The ideal estimators studied in [15] achieved bias reduction by adapting the bandwidth about each X_i to the size of $f(X_i)$, with smooth clipping for small values of $f(X_i)$, and using concentrated kernels, in order to keep the estimators local. Samiuddin & El-Sayyad [23] achieved the same results by shifting the centers of the windows by random quantities. Giné & Sang [10] studied the uniform convergence in almost sure sense of the true estimators corresponding to these ideal estimators.

1.3 KERNEL-TYPE REGRESSION ESTIMATOR

Suppose $(X_1,Y_1),...,(X_n,Y_n)$ are n pairs of *i.i.d.* observations, where Y_1 is a bounded random variable. $X_i,Y_i\in\mathbb{R}$, f is the probability density function of X_1, and r is the regression function defined as the conditional mean of Y:

$$r(t) = \mathbb{E}(Y|X=t). \tag{1.3.1}$$

We want to estimate $r(t)$ using $(X_1, Y_1), ..., (X_n, Y_n)$. One commonly used nonparametric regression estimator for $r(t)$ was introduced independently by Nadaraya [19] and Watson [28]. The so called Nadaraya-Watson estimator is defined as

$$\hat{r}(t; h_n) = \frac{\sum_{i=1}^n K\left(\frac{t-X_i}{h_n}\right) Y_i}{\sum_{i=1}^n K\left(\frac{t-X_i}{h_n}\right)} = \frac{\frac{1}{nh_n}\sum_{i=1}^n K\left(\frac{t-X_i}{h_n}\right) Y_i}{\frac{1}{nh_n}\sum_{i=1}^n K\left(\frac{t-X_i}{h_n}\right)} = \frac{\hat{g}(t; h_n)}{\hat{f}(t; h_n)}. \tag{1.3.2}$$

If $f(t)$ and $r(t)$ have bounded second order derivatives and the random variable Y is bounded, then the bias has order $O(h_n^2)$ and the variance of (1.3.2) has order $O((nh_n)^{-1})$. The proof of the bias of the Nadaraya-Watson estimator, (1.3.2), can be found in both [4] and [22]. Noda [20] established the convergence of the Nadaraya-Watson estimator to $r(t)$ and the mean square error at a continuous point. He also proved the consistency of estimator $\frac{\hat{g}(t;h_n)}{f(t)}$ almost surely when $f(t)$ known. The uniform consistency of the Nadaraya-Watson estimator was shown in [2] for the case of discrete $X_i's$. See [27] and [24] for more literature on the Nadaraya-Watson estimator.

Einmahl & Mason [6, 7] studied the estimator

$$\hat{r}_n(t, q) = \frac{\sum_{i=1}^n K\left(\frac{t-X_i}{h_n}\right) q(Y_i)}{nh_n \hat{f}(t, h_n)}, \tag{1.3.3}$$

where q is a chosen measurable function, for the regression function $r(t, q) := \mathbb{E}(q(Y)|X = t)$. Notice that for $q(Y) = Y$, we get the regression function (1.3.1). Einmahl & Mason [6] obtained the exact results for the rate of uniform consistency of regression estimators with some additional smoothness conditions for t in a compact interval. The rate of uniform consistency of the regression estimators was of order $O_{a.s.}(\sqrt{nh_n|\log h_n|})$.

1.4 VARIABLE BANDWIDTH KERNEL REGRESSION ESTIMATOR

Define $g(t) := r(t)f(t)$. Consider $\alpha(x)$ defined by (1.2.2) and then we have $\alpha(g(t))$ by

$$\alpha(g(t)) = \alpha(r(t)f(t)). \tag{1.4.1}$$

We propose and study the following version of the variable bandwidth kernel regression estimator:

$$\bar{r}(t;h_n) = \frac{\sum_{i=1}^{n} K\left(\frac{t-X_i}{h_n}\alpha(g(X_i))\right)\alpha(g(X_i))Y_i}{\sum_{i=1}^{n} K\left(\frac{t-X_i}{h_n}\alpha(f(X_i))\right)\alpha(f(X_i))} \tag{1.4.2}$$

$$= \frac{\frac{1}{nh_n}\sum_{i=1}^{n} K\left(\frac{t-X_i}{h_n}\alpha(g(X_i))\right)\alpha(g(X_i))Y_i}{\frac{1}{nh_n}\sum_{i=1}^{n} K\left(\frac{t-X_i}{h_n}\alpha(f(X_i))\right)\alpha(f(X_i))} = \frac{\bar{g}(t;h_n)}{\bar{f}(t;h_n)}. \tag{1.4.3}$$

One of the few studies on variable bandwidth kernel regression estimators was done by Müller & Stadtmüller [18]. They introduced the following variable bandwidth kernel estimator for regression curve:

$$\hat{r}(t;h_t) = \frac{1}{nh_t}\sum_{i=1}^{n} K\left(\frac{t-x_i}{h_t}\right)Y_i, \tag{1.4.4}$$

for a given function $t \to h_t$, $t \in [0,1]$. For $r(t) \in C^k[0,1]$ (we say that a function r is in $C^k[0,1]$ if it and its first k derivatives are bounded and uniformly continuous on $[0,1]$), the optimal bandwidth for h_t has the order of $O(n^{-1/(2k+1)})$. To estimate the regression function $r(t)$ in (1.3.1), Müller & Stadtmüller [18] worked on variation of bandwidth depending on t whereas we study bandwidth depending on the sample. Einmahl & Mason [7] also worked on establishing consistence of kernel-type estimators (1.3.3) in the multidimensional case when the bandwidth h_n is a function of the location t or the data. We will study the true variable bandwidth kernel regression estimator as an analog of the true variable density estimator:

$$\begin{aligned}
\hat{r}_n(t;h_{1,n},h_{2,n}) &= \frac{\sum_{i=1}^{n} K\left(\frac{t-X_i}{h_{2,n}}\alpha(\hat{g}(X_i;h_{1,n}))\right)\alpha(\hat{g}(X_i;h_{1,n}))Y_i}{\sum_{i=1}^{n} K\left(\frac{t-X_i}{h_{2,n}}\alpha(\hat{f}(X_i;h_{1,n}))\right)\alpha(\hat{f}(X_i;h_{1,n}))} \\
&= \frac{\frac{1}{nh_{2,n}}\sum_{i=1}^{n} K\left(\frac{t-X_i}{h_{2,n}}\alpha(\hat{g}(X_i;h_{1,n}))\right)\alpha(\hat{g}(X_i;h_{1,n}))Y_i}{\frac{1}{nh_{2,n}}\sum_{i=1}^{n} K\left(\frac{t-X_i}{h_{2,n}}\alpha(\hat{f}(X_i;h_{1,n}))\right)\alpha(\hat{f}(X_i;h_{1,n}))} \\
&= \frac{\hat{g}(t;h_{1,n},h_{2,n})}{\hat{f}(t;h_{1,n},h_{2,n})}, \qquad (1.4.5)
\end{aligned}$$

where $\hat{g}(t;h_{1,n}) = \hat{r}(t;h_{1,n})\hat{f}(t;h_{1,n})$.

1.5 OVERVIEW

1.5 Contribution of the dissertation

The contribution of this dissertation is as follows:

- We showed that the optimal bandwidth of the true variable bandwidth kernel density estimator is of the form $Jn^{-1/(8+d)}$, for some finite constant J.
- We have established central limit theorems for the ideal and true variable bandwidth kernel density estimators.
- We have extended the ideal and true variable bandwidth kernel density estimators to kernel regression.
- We have estimated the bias for both the ideal and true variable bandwidth kernel regression estimators.
- We have found the asymptotic variance for the ideal and true variable bandwidth kernel regression estimators.
- We have established central limit theorems for the ideal and true variable bandwidth kernel regression estimators.

1.5 Structure of the Dissertation

This dissertation is organized in the following way. In Section 2 of Chapter 2 and Section 3.1 of Chapter 3 we discuss the expansion of functions $K\left(\frac{t-X_i}{h_{2,n}}\alpha(\hat{f}(X_i;h_{1,n}))\right)$, $\alpha(\hat{f}(X_i;h_{1,n}))$, $K\left(\frac{t-X_i}{h_{2,n}}\alpha(\hat{g}(X_i;h_{1,n}))\right)$, and $\alpha(\hat{g}(X_i;h_{1,n}))$. We will also assess the convergence rate of the terms in the expansions. Section 2.2 focuses on the bias of the true variable bandwidth kernel density estimator in the multidimensional case. The variance of the ideal variable bandwidth kernel density estimator for the multidimensional case and the variance of the true variable bandwidth kernel density estimator for the one dimensional case are presented in Section 2.3. Section 2.4 provides central limit theorems for the ideal and true variable bandwidth kernel density estimators. Section 2.5 provides a simulation study and optimal bandwidth selection for the true variable bandwidth kernel density estimator. In Section 3.2 of Chapter 3, the bias of the ideal and true variable bandwidth kernel regression estimators for the one dimensional case is estimated. Section 3.3 presents central limit theorems for the ideal and true variable bandwidth kernel regression estimators. Chapter 4 gives the conclusion.

2 VARIABLE BANDWIDTH KERNEL DENSITY ESTIMATION

2.1 DECOMPOSITION FOR KERNEL DENSITY ESTIMATION

For convenience, we adopt the notations of [10] for the Taylor expansion of $K\left(\frac{t-X_i}{h_{2,n}}\alpha(\hat{f}(X_i;h_{1,n}))\right)$ at $K\left(\frac{t-X_i}{h_{2,n}}\alpha(f(X_i))\right)$. For details, see [10]. $\mathcal{P}_C$ denotes the set of all probability densities on $\mathbb{R}^d$ that are uniformly continuous and are bounded by $C < \infty$, and $\mathcal{P}_{C,k}$ denotes the set of densities on $\mathbb{R}^d$ where their partial derivatives of order k or lower are bounded by $C < \infty$ and are uniformly continuous. We say that a function g is in $C^l(\Omega)$ if it and its first l derivatives are bounded and uniformly continuous on Ω.

Define $\delta(t) = \delta(t, n)$ by the equation

$$\delta(t) = \frac{\alpha(\hat{f}(t;h_{1,n})) - \alpha(f(t))}{\alpha(f(t))}. \tag{2.1.1}$$

Then,

$$\alpha(\hat{f}(t;h_{1,n})) = \alpha(f(t))(1 + \delta(t)) \tag{2.1.2}$$

and

$$|\delta(t)| \leq Bc^{-2}|\hat{f}(t;h_{1,n}) - f(t)| \tag{2.1.3}$$

for a constant B that depends only on p and for $c > 0$. Although we study the asymptotics of the true estimator pointwise, the uniform asymptotic behavior of the quantity $\delta(\cdot)$ is needed in the latter analysis. Define

$$D_1(t;h_{1,n}) = \hat{f}(t;h_{1,n}) - \mathbb{E}\hat{f}(t;h_{1,n}) \;\text{ and }\; b_1(t;h_{1,n}) = \mathbb{E}\hat{f}(t;h_{1,n}) - f(t).$$

Note that for $f \in \mathcal{P}_{C,2}$ by (3.5) of [10],

$$\sup_{t \in \mathbb{R}^d} |b_1(t; h_{1,n})| = O(h_{1,n}^2), \tag{2.1.4}$$

and by [8],

$$\sup_{t \in \mathbb{R}^d} |D_1(t; h_{1,n})| = O_{a.s.}\left(\sqrt{\frac{\log n}{n h_{1,n}^d}}\right).$$

Denote

$$\sqrt{\frac{\log n}{n h_{1,n}^d}} + h_{1,n}^2 := U(h_{1,n}). \tag{2.1.5}$$

Then, we have by (3.6) of [10],

$$\sup_{t \in \mathbb{R}^d} |\delta(t)| = \sup_{t \in \mathbb{R}^d} |\hat{f}(t; h_{1,n}) - f(t)| = \sup_{t \in \mathbb{R}^d} |D_1(t; h_{1,n}) + b_1(t; h_{1,n})| = O_{a.s.}\left(U(h_{1,n})\right) \tag{2.1.6}$$

for $f \in \mathcal{P}_{C,2}$. By the definition of $\delta(t)$, we also have,

$$\delta(t) = \frac{\alpha'(f(t))[\hat{f}(t; h_{1,n}) - f(t)]}{\alpha(f(t))} + \frac{\alpha''(\eta)[\hat{f}(t; h_{1,n}) - f(t)]^2}{2\alpha(f(t))} \tag{2.1.7}$$

where $\eta = \eta(t, h_{1,n}) \geq 0$ is between $\hat{f}(t; h_{1,n})$ and $f(t)$. Note that, since $p \geq 1$ and when p' and p'' are uniformly bounded on $[0, \infty)$, we have $|\alpha''(\eta(t, h_{1,n}))| \leq c^{-3} A$ for some constant A which depends only on the clipping function p. It is also convenient to record the following expansion of $\alpha^d(\hat{f})$ implied by (2.1.2) and (2.1.6):

$$\alpha^d(\hat{f}(t; h_{1,n})) = \alpha^d(f(t))(1 + d\delta(t)) + \delta_1(t). \tag{2.1.8}$$

By (3.9) of [10], we have

$$\|\delta_1\|_\infty = O_{\text{a.s.}}(\|\delta\|_\infty^2) \quad \text{for} \quad f \in \mathcal{P}_{C,2}.$$

Hence, by (2.1.3) and (2.1.6),

$$\|\delta_1\|_\infty = O_{\text{a.s.}}(\|\hat{f}_n(\cdot; h_{1,n}) - f(\cdot)\|_\infty^2) \quad \text{for} \quad f \in \mathcal{P}_{C,2}.$$

Set

$$L_1(t) = \sum_{i=1}^{d} t_i K_i'(t) \quad \text{and} \quad L(t) = dK(t) + L_1(t), \quad t \in \mathbb{R}^d, \tag{2.1.9}$$

where K_i' denotes the partial derivative of K in the direction of the i-th coordinate, and t_i denotes the i-th coordinate of $t \in \mathbb{R}^d$. By symmetry and integration by parts, we notice that L is a second order kernel. We say that M is a second order kernel if $\int uM(u)du = 0$ and $\int u^2 M(u)du > 0$.

We then have the following Taylor expansion

$$K\left(\frac{t - X_i}{h_{2,n}}\alpha(\hat{f}(X_i; h_{1,n}))\right) = K\left(\frac{t - X_i}{h_{2,n}}\alpha(f(X_i))\right)$$
$$+ \sum_{j=1}^{d} K_j'\left(\frac{t - X_i}{h_{2,n}}\alpha(f(X_i))\right)\frac{(t - X_i)_j}{h_{2,n}}\alpha(f(X_i))\delta(X_i) + \delta_2(t; X_i), \tag{2.1.10}$$

where

$$\delta_2(t, X_i) = \sum_{j,\ell=1}^{d} K_{j,\ell}''(\xi)\frac{(t - X_i)_j(t - X_i)_\ell}{2h_{2,n}^2}\alpha^2(f(X_i))\delta^2(X_i), \tag{2.1.11}$$

ξ being a number between $\frac{t-X_i}{h_{2,n}}\alpha(f(X_i))$ and $\frac{t-X_i}{h_{2,n}}\alpha(f(X_i))(1+\delta(X_i))$. By the analysis (3.12) and (3.13) in [10],

$$\sup_{t,x\in\mathbb{R}^d} |\delta_2(t, x)| = O_{\text{a.s.}}\left(\|\hat{f}(\cdot; h_{1,n}) - f(\cdot)\|_\infty^2\right) = O_{a.s.}(U^2(h_{1,n})) \tag{2.1.12}$$

for $f \in \mathcal{P}_{C,2}$. Therefore, by using the expansion (2.1.8) of $\alpha^d(\hat{f})$, the Taylor expansion (2.1.10), and the notations L_1 and L in (2.1.9), we get

$$
\begin{aligned}
\hat{f}(t; h_{1,n}, h_{2,n}) = \bar{f}(t; h_{2,n}) & \\
&+ \frac{1}{nh_{2,n}^d} \sum_{i=1}^{n} L\left(\frac{t - X_i}{h_{2,n}} \alpha(f(X_i))\right) \alpha^d(f(X_i))\delta(X_i) \\
&+ \frac{1}{nh_{2,n}^d} \sum_{i=1}^{n} \Bigg[K\left(\frac{t - X_i}{h_{2,n}} \alpha(f(X_i))\right) \delta_1(X_i) + \alpha^d(f(X_i))\delta_2(t, X_i) \\
&\qquad + dL_1\left(\frac{t - X_i}{h_{2,n}} \alpha(f(X_i))\right) \alpha^d(f(X_i))\delta^2(X_i) \Bigg] \\
&+ \frac{1}{nh_{2,n}^d} \sum_{i=1}^{n} \Bigg[L_1\left(\frac{t - X_i}{h_{2,n}} \alpha(f(X_i))\right) \delta(X_i)\delta_1(X_i) \\
&\qquad + d\alpha^d(f(X_i))\delta(X_i)\delta_2(t, X_i) \Bigg] \\
&+ \frac{1}{nh_{2,n}^d} \sum_{i=1}^{n} \delta_2(t, X_i)\delta_1(X_i). \qquad (2.1.13)
\end{aligned}
$$

2.2 BIAS OF VARIABLE BANDWIDTH KERNEL DENSITY ESTIMATOR

2.2 Notations and Assumptions

Let $v = (v_1, \ldots, v_d) \in (\mathbb{N} \cup \{0\})^d$. $|v| = \sum_{i=1}^d v_i$, $D_v := D_{x_1}^{v_1} \circ \cdots \circ D_{x_d}^{v_d}$, $v! = v_1! \cdots v_d!$ and $\tau_v = \int_{\mathbb{R}^d} u_1^{v_1} \cdots u_d^{v_d} K(u)du$. Let the kernel K below the following assumptions which are required throughout the chapter.

Assumptions 1. *Suppose that the kernel K on $\mathbb{R}^d$ is non-negative and has the form $K(t) = \Phi(\|t\|^2)$ for some positive real-valued even function Φ with uniformly bounded second derivative and with support contained in $[0, T]$, $T < \infty$.*

We also give assumptions on clipping function p and density function f.

Assumptions 2. *Let the clipping function p as in (1.2.2) have at least fifth order derivative and satisfy the conditions (1.2.3) and (1.2.4). Assume $f \in C^4(\mathbb{R}^d)$ is a density function.*

Define

$$\mathcal{D}_r := \{t \in \mathbb{R}^d : f(t) > r > t_0 c^2, \|t\| < 1/r\}, \quad r > 0. \tag{2.2.1}$$

Here, c and t_0 are the constants that appear in the clipping function α in (1.2.2).

2.2 Ideal Estimator

Corollary 2.2.1. ([16, 17]) *Let $\alpha(f(t)) = cp^{1/2}(c^{-2}f(t))$ for some $c > 0$. Define $\bar{f}(t; h_n)$ by equation (1.2.1). Then, under Assumptions 2 for the clipping function p and density function f,*

$$E\bar{f}(t; h_n) = f(t) + \left(\sum_{|v|=4} \tau_v D_v(1/f)/v!\right) h_n^4 + o(h_n^4) = f(t) + O\left(h_n^4\right) \tag{2.2.2}$$

as $h_n \to 0$, uniformly on $\mathcal{D}_r$.

2.2 True Estimator

Proposition 2.2.2. *Let $\alpha(f(t)) = cp^{1/2}(c^{-2}f(t))$ for some $c > 0$, and define $\hat{f}(t; h_{1,n}, h_{2,n})$ by (1.2.8). Assume that $U(h_{1,n}) = o(h_{2,n}^2)$. Then, as $h_{2,n} \to 0$, for $t \in \mathcal{D}_r$ under Assumptions 1 for K and Assumptions 2 for p and f,*

$$\mathbb{E}(\hat{f}(t; h_{1,n}, h_{2,n})) - f(t) = \left(\sum_{|v|=4} \tau_v D_v(1/f)/v!\right) h_{2,n}^4 + o(h_{2,n}^4).$$

Comparing the Corollary 2.2.1 and the Proposition 2.2.2, the bias of the ideal estimator and the true estimator are of the same order. Now, to proof the Proposition 2.2.2 we require the following two lemmas. The proof of the first lemma can be found in (3.20) of [10]. This lemma basically gives the almost sure (a.s.) convergence rate of some terms in the decomposition (2.1.13) of $\hat{f}(t; h_{1,n}, h_{2,n})$.

Lemma 2.2.3. ([10]) *Define L and L_1 as in (2.1.9), δ, δ_1, and δ_2, as in (2.1.1), (2.1.8), and (2.1.11) respectively. Under the assumptions of Proposition 2.2.2,*

$$\frac{1}{nh_{2,n}^d}\sup_{t\in\mathbb{R}^d}\left|\sum_{i=1}^n\left[K\left(\frac{t-X_i}{h_{2,n}}\alpha(f(X_i))\right)\delta_1(X_i)+\alpha^d(f(X_i))\delta_2(t,X_i)\right.\right.$$
$$\left.\left.+dL_1\left(\frac{t-X_i}{h_{2,n}}\alpha(f(X_i))\right)\alpha^d(f(X_i))\delta^2(X_i)\right]\right|=O_{a.s.}(U^2(h_{1,n})),$$
$$\frac{1}{nh_{2,n}^d}\sup_{t\in\mathbb{R}^d}\left|\sum_{i=1}^n\left[L_1\left(\frac{t-X_i}{h_{2,n}}\alpha(f(X_i))\right)\delta(X_i)\delta_1(X_i)\right.\right.$$
$$\left.\left.+d\alpha^d(f(X_i))\delta(X_i)\delta_2(t,X_i)\right]\right|=O_{a.s.}(U^3(h_{1,n})),$$
$$\text{and }\frac{1}{nh_{2,n}^d}\sup_{t\in\mathbb{R}^d}\left|\sum_{i=1}^n\delta_2(t,X_i)\delta_1(X_i)\right|=O_{a.s.}(U^4(h_{1,n})).$$

Lemma 2.2.4. *Define L as in (2.1.9) and δ as in (2.1.1). Under the assumptions of Proposition 2.2.2,*

$$\frac{1}{h_{2,n}^d}\mathbb{E}\left(L\left(\frac{t-X_1}{h_{2,n}}\alpha(f(X_1))\right)\alpha^d(f(X_1))\delta(X_1)\right)=o(h_{2,n}^4)$$

Proof. Using the decomposition (2.1.7) of $\delta(t)$, and the decomposition of $\hat{f}-f$ into variance D_1 and bias b_1, we obtain

$$\frac{1}{h_{2,n}^d}\mathbb{E}\left(L\left(\frac{t-X_1}{h_{2,n}}\alpha(f(X_1))\right)\alpha^d(f(X_1))\delta(X_1)\right)$$
$$=\frac{1}{dh_{2,n}^d}\mathbb{E}\left[L\left(\frac{t-X_1}{h_{2,n}}\alpha(f(X_1))\right)(\alpha^d)'(f(X_1))D_1(X_1;h_{1,n})\right] \tag{2.2.3}$$
$$+\frac{1}{dh_{2,n}^d}\mathbb{E}\left[L\left(\frac{t-X_1}{h_{2,n}}\alpha(f(X_1))\right)(\alpha^d)'(f(X_1))b_1(X_1;h_{1,n})\right] \tag{2.2.4}$$
$$+\mathbb{E}\left[\frac{1}{2nh_{2,n}^d}\sum_{i=1}^n\left(L\left(\frac{t-X_i}{h_{2,n}}\alpha(f(X_i))\right)(\alpha^{d-1})(f(X_i))\right.\right.$$
$$\left.\left.\times(\alpha^d)''(\eta(X_i))[\hat{f}(X_i;h_{1,n})-f(X_i)]^2\right)\right]. \tag{2.2.5}$$

By (2.1.12) or (3.24) of [10] and the boundedness of $\alpha''(\eta)$ and L, we obtain,

$$|(2.2.5)| = O\left(U^2(h_{1,n})\right) = o(h_{2,n}^4). \tag{2.2.6}$$

By (3.26) of [10], we have

$$|(2.2.4)| = O(h_{1,n}^2 h_{2,n}^2) \text{ for } f \in \mathcal{P}_{C,4}. \tag{2.2.7}$$

To estimate the order of (2.2.3), we need a result from U-statistics. Let H be an integrable function of two i.i.d. random variables X and Y. The U-statistic is defined as

$$U_n(H) = \frac{1}{n(n-1)} \sum_{1 \le i \ne j \le n} H(X_i, X_j), \tag{2.2.8}$$

where the variables X_i are $i.i.d.$ copies of X. The second order Hoeffding projection of $H(X,Y)$ is

$$\pi_2(H)(X,Y) = H(X,Y) - \mathbb{E}_X H(X,Y) - \mathbb{E}_Y H(X,Y) + \mathbb{E}H, \tag{2.2.9}$$

where $\mathbb{E}_X$ denotes conditional expectation on X given Y. If we define

$$H_t(X,Y) := L\left(\frac{t-X}{h_{2,n}} \alpha(f(X))\right) (\alpha^d)'(f(X)) K\left(\frac{X-Y}{h_{1,n}}\right), \tag{2.2.10}$$

then we can decompose the following quantity into a diagonal term and a U-statistic term,

$$\begin{aligned} &\frac{1}{nh_{2,n}^d} \sum_{i=1}^n L\left(\frac{t-X_i}{h_{2,n}} \alpha(f(X_i))\right) (\alpha^d)'(f(X_i)) D_1(X_i; h_{1,n}) \\ &= \frac{1}{n^2 h_{1,n}^d h_{2,n}^d} \sum_{i=1}^n (H_t(X_i, X_i) - \mathbb{E}_Y H_t(X_i, Y)) \end{aligned} \tag{2.2.11}$$

$$+ \frac{n-1}{nh_{1,n}^d h_{2,n}^d} U_n \left(H_t - \mathbb{E}_Y H_t(.,Y) \right)$$

$$= \frac{1}{n^2 h_{1,n}^d h_{2,n}^d} \sum_{i=1}^{n} (H_t(X_i, X_i) - \mathbb{E}_Y H_t(X_i, Y)) \tag{2.2.12}$$

$$+ \frac{n-1}{nh_{1,n}^d h_{2,n}^d} U_n \left(\pi_2(H_t(\cdot, \cdot)) \right) \tag{2.2.13}$$

$$+ \frac{n-1}{n^2 h_{1,n}^d h_{2,n}^d} \sum_{i=1}^{n} (\mathbb{E}_X H_t(X, X_i) - \mathbb{E} H_t). \tag{2.2.14}$$

There are two empirical process terms (2.2.12) and (2.2.14), and a canonical U-statistic term (2.2.13). Obviously, (2.2.13) and (2.2.14) have mean zero since

$$\mathbb{E} U_n \left(\pi_2(H_t(\cdot, \cdot)) \right) = \mathbb{E}(\mathbb{E}_X H_t(X,Y) - \mathbb{E} H_t) = 0. \tag{2.2.15}$$

In the first empirical process (2.2.12), set $\bar{Q}_i(t) = H_t(X_i, X_i) - \mathbb{E}_Y H_t(X_i, Y)$ and by the boundedness of L, α, and f observe that with change of variable $u = (t-x)/h_{2,n}$,

$$\begin{aligned}
\mathbb{E}|\bar{Q}_1(t)| = \int_{\mathbb{R}^d} \Bigg| & L\left(\frac{t-x}{h_{2,n}} \alpha(f(x)) \right) (\alpha^d)'(f(x)) K(0) \\
& - \int_{\mathbb{R}^d} L\left(\frac{t-x}{h_{2,n}} \alpha(f(x)) \right) (\alpha^d)'(f(x)) K\left(\frac{x-y}{h_{1,n}} \right) dy \Bigg| dx \\
= h_{2,n}^d \int_{\mathbb{R}^d} \Bigg| & L\left(u\alpha(f(t-uh_{2,n})) \right) (\alpha^d)'(f(t-uh_{2,n})) K(0) \\
& - \int_{\mathbb{R}^d} L\left(u\alpha(f(t-uh_{2,n})) \right) (\alpha^d)'(f(t-uh_{2,n})) K\left(\frac{t-uh_{2,n}-y}{h_{1,n}} \right) dy \Bigg| du \\
\le B h_{2,n}^d, &
\end{aligned} \tag{2.2.16}$$

for some finite constant B. By combining the above analysis for $\mathbb{E}|\bar{Q}_1(t)|$ and (2.2.15), we have

$$
\begin{aligned}
(2.2.3) &= \frac{1}{d}\mathbb{E}((2.2.11)) \\
&= \frac{1}{d}\mathbb{E}\left(\frac{1}{n^2 h_{1,n}^d h_{2,n}^d}\sum_{i=1}^{n}(H_t(X_i,X_i) - \mathbb{E}_Y H_t(X_i,Y))\right) \\
&= O\left(\frac{1}{nh_{1,n}^d}\right). \qquad (2.2.17)
\end{aligned}
$$

Thus, by the formulas (2.2.6), (2.2.7), (2.2.17), and under the condition that $U(h_{1,n}) = o(h_{2,n}^2)$, we have

$$
\frac{1}{h_{2,n}^d}\mathbb{E}\left(L\left(\frac{t-X_1}{h_{2,n}}\alpha(f(X_1))\right)\alpha^d(f(X_1))\delta(X_1)\right) = o(h_{2,n}^4)
$$

$\square$

Proof of Proposition 2.2.2. By decomposition (2.1.13) of $\hat{f}(t; h_{1,n}, h_{2,n})$, we obtain

$$
\begin{aligned}
&\mathbb{E}\hat{f}(t; h_{1,n}, h_{2,n}) = \mathbb{E}\bar{f}(t; h_{2,n}) &(2.2.18)\\
&+ \frac{1}{h_{2,n}^d}\mathbb{E}\left(L\left(\frac{t-X_1}{h_{2,n}}\alpha(f(X_1))\right)\alpha^d(f(X_1))\delta(X_1)\right) &(2.2.19)\\
&+ \mathbb{E}\left[\frac{1}{nh_{2,n}^d}\sum_{i=1}^{n}\left(K\left(\frac{t-X_i}{h_{2,n}}\alpha(f(X_i))\right)\delta_1(X_i) + \alpha^d(f(X_i))\delta_2(t,X_i)\right.\right. \\
&\left.\left. + dL_1\left(\frac{t-X_i}{h_{2,n}}\alpha(f(X_i))\right)\alpha^d(f(X_i))\delta^2(X_i)\right)\right] &(2.2.20)\\
&+ \mathbb{E}\left[\frac{1}{nh_{2,n}^d}\sum_{i=1}^{n}\left(L_1\left(\frac{t-X_i}{h_{2,n}}\alpha(f(X_i))\right)\delta(X_i)\delta_1(X_i) + d\alpha^d(f(X_i))\delta(X_i)\delta_2(t,X_i)\right)\right] \\
& &(2.2.21)\\
&+ \mathbb{E}\left[\frac{1}{nh_{2,n}^d}\sum_{i=1}^{n}\delta_2(t,X_i)\delta_1(X_i)\right]. &(2.2.22)
\end{aligned}
$$

We show that the terms (2.2.19)-(2.2.22) is of order $o(h_{2,n}^4)$. By Lemma 2.2.3 and the boundedness of α, K and L_1, and since $U(h_{1,n}) = o(h_{2,n}^2)$, we have

$$|(2.2.20)| = O\left(U^2(h_{1,n})\right) = o(h_{2,n}^4), \tag{2.2.23}$$

$$|(2.2.21)| = O\left(U^3(h_{1,n})\right) = o(h_{2,n}^6), \tag{2.2.24}$$

$$\text{and } |(2.2.22)| = O\left(U^4(h_{1,n})\right) = o(h_{2,n}^8). \tag{2.2.25}$$

By [16, 17] or Corollary 1 of [10], the ideal estimator (1.2.1) satisfies

$$\mathbb{E}\bar{f}(t; h_{2,n}) = f(t) + \left(\sum_{|v|=4} \tau_v D_v(1/f)/v!\right) h_{2,n}^4 + o(h_{2,n}^4). \tag{2.2.26}$$

By the analysis in (2.2.23)-(2.2.25), and Lemma 2.2.4, we have that the order of the terms (2.2.19) -(2.2.22) in $\mathbb{E}(\hat{f}(t; h_{1,n}, h_{2,n}))$ is of order $o(h_{2,n}^4)$.

Hence, by above analysis and (2.2.26), we have

$$\mathbb{E}(\hat{f}(t; h_{1,n}, h_{2,n})) - f(t) = \left(\sum_{|v|=4} \tau_v D_v(1/f)/v!\right) h_{2,n}^4 + o(h_{2,n}^4),$$

when $U(h_{1,n}) = o(h_{2,n}^2)$ and by the boundedness of K, L_1, and α. □

2.3 VARIANCE OF VARIABLE BANDWIDTH KERNEL DENSITY ESTIMATOR

2.3 Ideal Estimator

In this subsection, we present the results on the variance of the ideal estimator $\bar{f}(t; h_n)$ with the following propositions.

Proposition 2.3.1. *Let* $\alpha(f(t)) = cp^{1/2}(c^{-2}f(t))$ *for some* $c > 0$, *and* p *in Assumptions 1. Define* $\bar{f}(t;h_n)$ *by equation (1.2.1). Then, under Assumptions 1,*

$$Var(\bar{f}(t;h_n)) = \frac{\alpha^d(f(t))f(t)\mu_0}{nh^d}(1+o(1))$$

as $h_n \to 0$, *uniformly on* $\mathcal{D}_r$ *where* $\mu_0 = \int_{\mathbb{R}^d} K^2(w)dw$.

The next proposition is necessary to get the expansion of the quantity $\mathbb{E}A^2(X_1)$, where $A(X_i) = \gamma^d(X_i)K(h^{-1}\gamma(X_i)(t - X_i))$. The idea is similar to the uniform bias expansion of [17] (Theorems 1.1, 2.10, and 5.13), [15] (Theorem A.1), and particularly [10]. So, we sketch McKay [17]'s proof for the case $d = 1$.

Proposition 2.3.2. *Let* K *on* $\mathbb{R}^d$ *be symmetric about zero and have bounded support. Assume that* $\zeta \in C^l(\mathbb{R}^d)$. *Suppose that* $\gamma(t) \geq c > 0$ $\forall t \in \mathbb{R}^d$, *and* $\gamma(t) \in C^{l+1}(\mathbb{R}^d)$. *Then,*

$$\int \gamma^{2d}(s)\zeta(s)K^2\left(\frac{t-s}{h}\gamma(s)\right)ds = \sum_{k=0}^{l} a_k(t)h^{k+d} + o(h^{l+d}) \tag{2.3.1}$$

as $h \to 0$, *uniformly in* t. *The functions* $a_k(t)$ *are uniformly bounded and equicontinuous and are defined for* $k \leq l/2$ *by*

$$a_{2k+1}(t) = 0, \quad a_{2k}(t) = \sum_{|v|=2k} \frac{\mu_v}{v!} D_v\left(\frac{\zeta(t)}{\gamma^{2k-d}(t)}\right). \tag{2.3.2}$$

In particular, $a_0(t) = \gamma^d(t)\zeta(t)\mu_0$, $\zeta \in C^l(\mathbb{R}^d)$ *where* $\mu_v = \int w_1^{v_1} \cdots w_d^{v_d} K^2(w)dw$.

Proof. (For $d = 1$.) Note that there exists $\epsilon > 0$ such that $\gamma(t-v) - v\gamma'(t-v)$ is bounded away from zero for all $t \in \mathbb{R}$ and $v \in [-\epsilon, \epsilon]$ for functions γ that are bounded away from zero and their derivatives that are bounded. Thus, $U_t(v) = v\gamma(t-v)$ is an invertible function on $[-\epsilon, \epsilon]$. By differentiation, we have that this inverse function, denoted by $V_t(u)$, are $l+1$ times differentiable with continuous derivatives. Unless $|t-s| \leq hT/c$, $K(h^{-1}\gamma(s)(t-s)) = 0$

for K in $[-T, T]$. Hence, the change of variable

$$hz = U_t(t-s), t-s = V_t(hz),$$

in the following integral is valid for all h small enough

$$\begin{aligned} & \int \gamma^2(s)\zeta(s)K^2\left(\frac{t-s}{h}\gamma(s)\right)ds \qquad (2.3.3) \\ = -h & \int \gamma^2(t-V_t(hz))\zeta(t-V_t(hz))\frac{dV_t(hz)}{d(hz)}K^2(z)dz. \end{aligned}$$

If we develop the function $\gamma^2(t-V_t(hz))\zeta(t-V_t(hz))\frac{dV_t(hz)}{d(hz)}$ into powers of hz and then integrate it, noting the compactness of the domain of integration ($z \in [-T, T]$) and the differentiability properties of ζ and γ, we have (2.3.1).

Let ψ be infinitely differentiable and has bounded support. First, changing variables ($t = s + hu$) from t to u in (2.3.3), we get

$$\begin{aligned} & \int \psi(t)(2.3.3)dt = h\int\int \psi(s+hu)\gamma^2(s)\zeta(s)K^2(u\gamma(s))duds \\ = \quad & h\int \gamma^2(s)\zeta(s)\int \psi(s+hu)K^2(u\gamma(s))duds. \end{aligned}$$

Next, developing ψ, changing variables once more ($w = u\gamma(s))$), we get

$$\int \psi(t)(2.3.3)dt = h\int \gamma^2(s)\zeta(s)\int \sum_{k=0}^{l} \frac{\psi^{(k)}(s)}{k!}h^k u^k K^2(u\gamma(s))duds + o(h^{l+1})$$

$$\begin{aligned} = \quad & h\int \gamma(s)\zeta(s)\int \sum_{k=0}^{l} \frac{\psi^{(k)}(s)}{k!}\frac{h^k w^k}{\gamma^k(s)}K^2(w)dwds + o(h^{l+1}) \\ = \quad & \sum_{k=0}^{l} \mu_k h^{k+1}\int \zeta(s)\frac{\psi^{(k)}(s)}{k!\gamma^{k-1}(s)}ds + o(h^{l+1}). \end{aligned}$$

Finally, by integrating by parts, we obtain

$$\int \psi(t)(2.3.3)dt = \int \sum_{k=0}^{l} \mu_k h^{k+1} k!^{-1} \psi(s) \frac{d^k(\zeta(s)\gamma^{1-k}(s))}{ds^k} ds + o(h^{l+1}). \tag{2.3.4}$$

Notice that $\mu_{2k+1} = 0$ for $k \geq 0$. Then, (2.3.2) follows by comparing the coefficients of h^k in both expansions (2.3.1) and (2.3.4). □

Remark 2.3.5. *For the d dimensional case, the change of variables in (2.3.3) and (2.3.4) gives h^d instead of h. By the results from [17] (Theorems 1.1, 2.10, and 5.13) for the multidimensional case we can verify the result in (2.3.1).*

Proof of Proposition 2.3.1. We will develop the second moment expansion uniformly to deal with the variance of the ideal estimator. Here we denote $h = h_n$ for convenience. The ideal estimator (1.2.1) has the form

$$\bar{f}(t;h) = \frac{1}{nh^d} \sum_{i=1}^{n} \alpha^d(f(X_i)) K(h^{-1}\alpha(f(X_i))(t - X_i)), \quad t \in \mathbb{R}^d.$$

The second moment of the ideal estimator is

$$\begin{aligned}
\mathbb{E}\bar{f}^2(t;h) &= \frac{1}{n^2h^{2d}} \mathbb{E} \left(\sum_{i=1}^{n} A(X_i) \right)^2 \\
&= \frac{1}{n^2h^{2d}} \sum_{i=1}^{n} \mathbb{E}A^2(X_i) + \frac{1}{n^2h^{2d}} \sum_{i \neq j} \mathbb{E}A(X_i)\mathbb{E}A(X_j) \\
&= \frac{1}{nh^{2d}} \mathbb{E}A^2(X_1) + \frac{n(n-1)}{n^2h^{2d}} (\mathbb{E}A(X_1))^2.
\end{aligned}$$

Hence, we obtain

$$\begin{aligned} Var\bar{f}(t;h) &= \mathbb{E}\bar{f}^2(t;h) - [\mathbb{E}\bar{f}(t;h)]^2 \\ &= \frac{1}{nh^{2d}}\mathbb{E}A^2(X_1) + \frac{(n-1)}{nh^{2d}}(\mathbb{E}A(X_1))^2 - \frac{1}{h^{2d}}(\mathbb{E}A(X_1))^2 \\ &= \frac{1}{nh^{2d}}\mathbb{E}A^2(X_1) + \frac{1}{nh^{2d}}(\mathbb{E}A(X_1))^2. \end{aligned}$$

By Proposition 1 of [10] or Corollary 2.2.1, we have

$$\mathbb{E}\bar{f}(t;h) = \frac{1}{h^d}\mathbb{E}A(X_1) = f(t) + O(h^2). \tag{2.3.6}$$

Then, by (2.3.6), we obtain

$$\begin{aligned} Var\bar{f}(t;h) &= \frac{1}{nh^{2d}}\mathbb{E}A^2(X_1) - \frac{1}{n}\left[f(t) + O(h^2)\right]^2 \\ &= \frac{1}{nh^{2d}}\mathbb{E}A^2(X_1) + O(n^{-1}). \end{aligned} \tag{2.3.7}$$

By Proposition 2.3.2 using $\zeta(x) = f(x)$ and $\gamma(x) = \alpha(f(x))$, we have

$$\mathbb{E}A^2(X_1) = h^d\alpha^d(f(t))f(t)\mu_0 + o(h^d). \tag{2.3.8}$$

Thus, the variance of the ideal estimator is $\frac{\alpha^d(f(t))f(t)\mu_0}{nh^d}(1+o(1))$ by applying Proposition 2.3.2, (2.3.8) and (2.3.7).

□

2.3 True Estimator in One Dimension

In this subsection, we present the results on the variance of the true estimator $\hat{f}(t;h_{1,n},h_{2,n})$.

We denote $K_t(x) = K\left(\frac{t-x}{h_{2,n}}\alpha(f(x))\right)$, $L_t(x) = L\left(\frac{t-x}{h_{2,n}}\alpha(f(x))\right)$, $L_{t,1}(x) = L_1\left(\frac{t-x}{h_{2,n}}\alpha(f(x))\right)$, $M(x) = L_t(x)\alpha'(f(x))$ and $M_i = M(X_i)$.

Theorem 2.3.3. *Let $X_1, ..., X_n$ be a random sample of size n with density function $f(t)$, $t \in \mathbb{R}$. Consider $\hat{f}(t; h_{1,n}, h_{2,n})$ defined in* (1.2.8). *Assume that the kernel K is a symmetric function with support contained in $[-T, T]$, $T < \infty$, and has bounded second order derivatives. The function $\alpha(x)$ is defined in (1.2.2) for a nondecreasing clipping function $p(s)$. Suppose that $U(h_{1,n}) = o(h_{2,n}^2)$ and $h_{2,n} = n^{-1/9}$. Then, under Assumptions 2 for p and f,*

$$Var(\hat{f}(t; h_{1,n}, h_{2,n})) = (V_1 + V_2 + V_3)(nh_{2,n})^{-1}(1 + o(1)) \tag{2.3.9}$$

where

$$V_1 = (h_{2,n})^{-1}\mathbb{E}\left(K_t^2(X_1)\alpha^2(f(X_1))\right) = \alpha(f(t))f(t)\mu_0(1 + o(1)),$$

$$V_2 = \frac{2}{h_{1,n}h_{2,n}}\mathbb{E}\left[K_t(X_2)\alpha(f(X_2))M_1 K\left(\frac{X_1 - X_2}{h_{1,n}}\right)\right],$$

and

$$V_3 = \frac{1}{h_{1,n}^2 h_{2,n}}\mathbb{E}\left[M_1 M_2 K\left(\frac{X_1 - X_3}{h_{1,n}}\right) K\left(\frac{X_2 - X_3}{h_{1,n}}\right)\right]. \tag{2.3.10}$$

Proof. By definition,

$$Var(\hat{f}(t; h_{1,n}, h_{2,n})) = \mathbb{E}\hat{f}^2(t; h_{1,n}, h_{2,n}) - (\mathbb{E}\hat{f}(t; h_{1,n}, h_{2,n}))^2.$$

We shall evaluate $\mathbb{E}\hat{f}^2(t; h_{1,n}, h_{2,n})$.

$$\begin{aligned}
&\mathbb{E}\hat{f}^2(t; h_{1,n}, h_{2,n}) \\
&= \frac{1}{nh_{2,n}^2}\mathbb{E}\left(K^2\left(\frac{t - X_1}{h_{2,n}}\alpha(\hat{f}(X_1; h_{1,n}))\right)\alpha^2(\hat{f}(X_1; h_{1,n}))\right) &(2.3.11)\\
&+ \frac{n-1}{nh_{2,n}^2}\mathbb{E}\left(K\left(\frac{t - X_1}{h_{2,n}}\alpha(\hat{f}(X_1; h_{1,n}))\right)\right. \\
&\left.\times\alpha(\hat{f}(X_1; h_{1,n}))K\left(\frac{t - X_2}{h_{2,n}}\alpha(\hat{f}(X_2; h_{1,n}))\right)\alpha(\hat{f}(X_2; h_{1,n}))\right). &(2.3.12)
\end{aligned}$$

The decompositions of $\alpha(\hat{f}(X_1;h_{1,n}))$ in (2.1.2) and $K\left(\frac{t-X_1}{h_{2,n}}\alpha(\hat{f}(X_1;h_{1,n}))\right)$ in (2.1.10) then give:

$$(2.3.11) = (nh_{2,n}^2)^{-1}\mathbb{E}\left(K_t^2(X_1)\alpha^2(f(X_1))(1+\delta(X_1))^2\right) \tag{2.3.13}$$

$$+ (nh_{2,n}^2)^{-1}\mathbb{E}\left(L_{t,1}^2(X_1)\alpha^2(f(X_1))\delta^2(X_1)(1+\delta(X_1))^2\right) \tag{2.3.14}$$

$$+ (nh_{2,n}^2)^{-1}\mathbb{E}\left(\delta_2^2(t,X_1)\alpha^2(f(X_1))(1+\delta(X_1))^2\right)$$

$$+ 2(nh_{2,n}^2)^{-1}\mathbb{E}\left(K_t(X_1)\alpha^2(f(X_1))(1+\delta(X_1))^2 L_{t,1}(X_1)\delta(X_1)\right) \tag{2.3.15}$$

$$+ 2(nh_{2,n}^2)^{-1}\mathbb{E}\left(L_{t,1}(X_1)\delta(X_1)\delta_2(t,X_1)\alpha^2(f(X_1))(1+\delta(X_1))^2\right)$$

$$+ 2(nh_{2,n}^2)^{-1}\mathbb{E}\left(K_t(X_1)\delta_2(t,X_1)\alpha^2(f(X_1))(1+\delta(X_1))^2\right). \tag{2.3.16}$$

By (2.3.8) and since $f \in C^4(\mathbb{R})$, $\alpha \in C^5(\mathbb{R})$, and by bounded support and symmetric property of K, we have (2.3.13) $= \frac{V_1(1+o(1))}{nh_{2,n}} = \frac{\alpha(f(t))f(t)\mu_0}{nh_{2,n}}(1+o(1))$. Notice that each term in (2.3.14)-(2.3.16) contains multiple of either $\delta(\cdot)$ or $\delta_2(\cdot)$, which makes them decay as $o(n^{-1})$. Let us verify this only for the term in 2.3.15 , since the proof for the remaining terms follow similarly. When $U(h_{1,n}) = o(h_{2,n}^2)$ and $h_{2,n} = O(n^{-1/9})$, by the boundedness of K, L_1, $\alpha(f)$ (due to the boundedness of f) and (2.1.6) and (2.1.12), we have that

$$(2.3.15) = O\left(\frac{U(h_{1,n})}{nh_{2,n}^2}\right) = o(n^{-1}).$$

Next we will work on the cross product term (2.3.12). Again, by applying the decompositions of $\alpha(\hat{f}(X_1;h_{1,n}))$ in (2.1.2) and $K\left(\frac{t-X_i}{h_{2,n}}\alpha(\hat{f}(X_i;h_{1,n}))\right)$ in (2.1.10), and noting that the expectation terms with coefficients $-\frac{1}{nh_{2,n}^2}$ and the terms with coefficients $\frac{n}{nh_{2,n}^2} = \frac{1}{h_{2,n}^2}$ along with quantities $\delta^2(\cdot)\delta(\cdot)$, $\delta_2(\cdot)\delta(\cdot)$ or $\delta_2(\cdot)\delta_2(\cdot)$ are negligible compared to $O(nh_{2,n})^{-1}$, we get

$$(2.3.12) = (n-1)(nh_{2,n}^2)^{-1}\mathbb{E}\left[K_t(X_1)K_t(X_2)\alpha(f(X_1))\alpha(f(X_2))\right] \tag{2.3.17}$$

$$+ h_{2,n}^{-2}\mathbb{E}\left[L_t(X_1)L_t(X_2)\delta(X_1)\delta(X_2)\alpha(f(X_1))\alpha(f(X_2))\right] \tag{2.3.18}$$

$$+ 2h_{2,n}^{-2}\mathbb{E}\left[K_t(X_1)\alpha(f(X_1))L_t(X_2)\delta(X_2)\alpha(f(X_2))\right] \tag{2.3.19}$$

$$+ 2h_{2,n}^{-2}\mathbb{E}\left[K_t(X_1)\alpha(f(X_1))L_{t,1}(X_2)\delta^2(X_2)\alpha(f(X_2))\right] \tag{2.3.20}$$

$$+ 2h_{2,n}^{-2}\mathbb{E}\left[K_t(X_1)\alpha(f(X_1))\delta_2(t,X_2)\alpha(f(X_2))\right] + o((nh_{2,n})^{-1}). \tag{2.3.21}$$

On the other hand, by the boundedness of K, L_1, $\alpha(f)$ and (2.1.6) and (2.1.12), the terms in expansion of $\left(\mathbb{E}(\hat{f}(t;h_{1,n},h_{2,n}))\right)^2$ that include the quantities $\delta^2(\cdot)\delta(\cdot)$, $\delta_2(\cdot)\delta(\cdot)$ or $\delta_2(\cdot)\delta_2(\cdot)$ have order $o(\frac{1}{nh_{2,n}})$. Therefore,

$$\left(\mathbb{E}(\hat{f}(t;h_{1,n},h_{2,n}))\right)^2 = h_{2,n}^{-2}\left\{\mathbb{E}\left[K_t(X_1)\alpha(f(X_1))\right]\right\}^2 \tag{2.3.22}$$

$$+ h_{2,n}^{-2}\left[\mathbb{E}\left(L_t(X_1)\delta(X_1)\alpha(f(X_1))\right)\right]^2 \tag{2.3.23}$$

$$+ 2h_{2,n}^{-2}\mathbb{E}[K_t(X_1)\alpha(f(X_1)]\mathbb{E}[L_t(X_1)\delta(X_1)\alpha(f(X_1))] \tag{2.3.24}$$

$$+ 2h_{2,n}^{-2}\mathbb{E}[K_t(X_1)\alpha(f(X_1)]\mathbb{E}[L_{t,1}(X_1)\delta^2(X_1)\alpha(f(X_1))] \tag{2.3.25}$$

$$+ 2h_{2,n}^{-2}\mathbb{E}[K_t(X_1)\alpha(f(X_1))]\mathbb{E}[\delta_2(t,X_1)\alpha(f(X_1))] + o(nh_{2,n})^{-1}. \tag{2.3.26}$$

By the above analysis, we shall study the difference between the terms (2.3.13), (2.3.17)-(2.3.21) and (2.3.22)-(2.3.26) to get (2.3.9).

The difference between (2.3.17) and (2.3.22)

The difference between (2.3.17) and (2.3.22) is $n^{-1}\left[\mathbb{E}\left((h_{2,n})^{-1}K_t(X_1)\alpha(f(X_1))\right)\right]^2$, which has order $O(n^{-1})$ due to the boundedness of K, α and f and by the bias formula of the ideal estimator from Proposition 1 of [10].

The difference between (2.3.18) and (2.3.23)

Since L has bounded support, $\alpha(f(X))$ is bounded and bounded away from zero and with change of variable $u = (t-x)/h_{2,n}$,

$$\begin{aligned}\mathbb{E}|M_1| &\leq \int L\left(\frac{t-x}{h_{2,n}}\alpha(f(x))\right)|\alpha'(f(x))|f(x)dx \\ &= h_{2,n}\int L\left(u\alpha(f(t-uh_{2,n}))\right)|\alpha'(f(t-uh_{2,n}))|f(t-uh_{2,n})du = O(h_{2,n}).\end{aligned} \tag{2.3.27}$$

Recall the decomposition of $\delta(t)$ in (2.1.7). (2.1.4), (2.1.6) and (2.3.27) give

$$
\begin{aligned}
&h_{2,n}^{-1}\mathbb{E}\left(M_1 b_1(X_1;h_{1,n})\right) = O(h_{1,n}^2),\\
&h_{2,n}^{-1}\mathbb{E}\left(M_1[\hat{f}(X_1;h_{1,n}) - f(t)]\right) = O(U(h_{1,n})),\\
&h_{2,n}^{-1}\mathbb{E}\left(M_1\alpha''(\eta)[\hat{f}(X_1;h_{1,n}) - f(t)]^2\right) = O(U^2(h_{1,n})).
\end{aligned}
\tag{2.3.28}
$$

We first apply the decomposition (2.1.7) of $\delta(t)$, and use the above analysis. Then, by applying the results from section 3.1 and 3.2 of [10] and (2.3.28), we get the following,

$$
(2.3.23) = \left[h_{2,n}^{-1}\mathbb{E}\left(M_1 D_1(X_1;h_{1,n})\right) + h_{2,n}^{-1}\mathbb{E}\left(M_1 b_1(X_1;h_{1,n})\right)\right]^2 + O(U^3(h_{1,n})).
$$

Now, by (2.2.17) and since $h_{2,n}^{-1}\mathbb{E}\left(M_1 b_1(X_1;h_{1,n})\right) = O(h_{1,n}^2)$,

$$
(2.3.23) = \left[h_{2,n}^{-1}\mathbb{E}\left(M_1 b_1(X_1;h_{1,n})\right)\right]^2 + O\left(U^3(h_{1,n}) + h_{1,n}n^{-1} + n^{-2}h_{1,n}^{-2}\right). \tag{2.3.29}
$$

On the other hand, by a similar analysis,

$$
\begin{aligned}
(2.3.18) &= h_{2,n}^{-2}\mathbb{E}\left(M_1 M_2[D_1(X_1;h_{1,n}) + b_1(X_1;h_{1,n})]\right.\\
&\times \left.[D_1(X_2;h_{1,n}) + b_1(X_2;h_{1,n})]\right) + O(U^3(h_{1,n}))\\
&:= Q_1 + Q_2 + Q_3 + O(U^3(h_{1,n})),
\end{aligned}
$$

where

$$
\begin{aligned}
Q_1 &= h_{2,n}^{-2}\mathbb{E}(M_1 M_2 D_1(X_1;h_{1,n}) D_1(X_2;h_{1,n}))\\
Q_2 &= h_{2,n}^{-2}\mathbb{E}(M_1 M_2 D_1(X_1;h_{1,n}) b_1(X_2;h_{1,n}))\\
Q_3 &= h_{2,n}^{-2}\mathbb{E}(M_1 M_2 b_1(X_1;h_{1,n}) b_1(X_2;h_{1,n})) = h_{2,n}^{-2}[\mathbb{E}(M_1 b_1(X_1;h_{1,n}))]^2.
\end{aligned}
$$

We also denote $N_i = \int K\left(\frac{X_i-u}{h_{1,n}}\right) f(u)du, i = 1,2$. Then,

$$
\begin{aligned}
Q_1 = & \frac{1}{n^2h_{1,n}^2h_{2,n}^2}\mathbb{E}\left[M_1M_2(K(0)-N_1)(K(0)-N_2)\right] && (2.3.30)\\
+ & \frac{1}{n^2h_{1,n}^2h_{2,n}^2}\mathbb{E}\left[M_1M_2\left(K\left(\frac{X_1-X_2}{h_{1,n}}\right)-N_1\right)\left(K\left(\frac{X_2-X_1}{h_{1,n}}\right)-N_2\right)\right] && (2.3.31)\\
+ & \frac{2(n-2)}{n^2h_{1,n}^2h_{2,n}^2}\mathbb{E}\left[M_1M_2\left(K\left(\frac{X_1-X_2}{h_{1,n}}\right)-N_1\right)\left(K\left(\frac{X_2-X_3}{h_{1,n}}\right)-N_2\right)\right] && (2.3.32)\\
+ & \frac{n-2}{n^2h_{1,n}^2h_{2,n}^2}\mathbb{E}\left[M_1M_2\left(K\left(\frac{X_1-X_3}{h_{1,n}}\right)-N_1\right)\left(K\left(\frac{X_2-X_3}{h_{1,n}}\right)-N_2\right)\right] && (2.3.33)\\
+ & \frac{(n-2)(n-3)}{n^2h_{1,n}^2h_{2,n}^2}\mathbb{E}\left[M_1M_2\left(K\left(\frac{X_1-X_3}{h_{1,n}}\right)-N_1\right)\left(K\left(\frac{X_2-X_4}{h_{1,n}}\right)-N_2\right)\right] && (2.3.34)\\
+ & \frac{2}{n^2h_{1,n}^2h_{2,n}^2}\mathbb{E}\left[M_1M_2\left(K(0)-N_1\right)\left(K\left(\frac{X_2-X_1}{h_{1,n}}\right)-N_2\right)\right] && (2.3.35)\\
+ & \frac{2(n-2)}{n^2h_{1,n}^2h_{2,n}^2}\mathbb{E}\left[M_1M_2\left(K(0)-N_1\right)\left(K\left(\frac{X_2-X_3}{h_{1,n}}\right)-N_2\right)\right]. && (2.3.36)
\end{aligned}
$$

We have that (2.3.30), (2.3.31), (2.3.35) have the order of $O(n^{-2}h_{1,n}^{-2})$ by applying (2.3.27). Since

$$\mathbb{E}\left(K\left(\frac{X_2-X_3}{h_{1,n}}\right)|X_2\right) = \int K\left(\frac{X_2-u}{h_{1,n}}\right) f(u)du = N_2, \quad (2.3.37)$$

(2.3.32) $= 0$ by the independence property of X_i and law of total expectation $[\mathbb{E}(X) = \mathbb{E}(\mathbb{E}(X|Y))]$. Similarly, (2.3.34) $=$ (2.3.36) $= 0$.

From (3.37) of [10], we have $V_3 = O(1)$. One can also show that $\mathbb{E}(M_1N_1) = O(h_{1,n}h_{2,n})$. Hence,

$$
\begin{aligned}
(2.3.33) = & \frac{n-2}{n^2h_{1,n}^2h_{2,n}^2}\mathbb{E}\left[M_1M_2K\left(\frac{X_1-X_3}{h_{1,n}}\right)K\left(\frac{X_2-X_3}{h_{1,n}}\right)\right]\\
& -\frac{n-2}{n^2h_{1,n}^2h_{2,n}^2}[\mathbb{E}(M_1N_1)]^2 = \frac{V_3}{nh_{2,n}} + O(n^{-1})
\end{aligned}
$$

and $Q_1 = \frac{V_3}{nh_{2,n}} + O(n^{-1})$.

Now, to show that $Q_2 = O(h_{1,n}n^{-1})$, we show that the following decomposition of Q_2 is of the same order, i.e., $O(h_{1,n}n^{-1})$.

$$Q_2 = \frac{2}{h_{2,n}^2}\mathbb{E}\left(M_1M_2D_1(X_1;h_{1,n})b_1(X_2;h_{1,n})\right)$$

$$= \frac{2}{nh_{1,n}^2h_{2,n}^2}\mathbb{E}\left(M_1M_2\left[K\left(\frac{X_1-X_2}{h_{1,n}}\right)-N_1\right]\left[\int K\left(\frac{X_2-v}{h_{1,n}}\right)(f(v)-f(X_2))dv\right]\right) \tag{2.3.38}$$

$$+ \frac{2(n-2)}{nh_{1,n}^2h_{2,n}^2}\mathbb{E}\left(M_1M_2\left[K\left(\frac{X_1-X_3}{h_{1,n}}\right)-N_1\right]\left[\int K\left(\frac{X_2-v}{h_{1,n}}\right)(f(v)-f(X_2))dv\right]\right) \tag{2.3.39}$$

$$+ \frac{2}{nh_{1,n}^2h_{2,n}^2}\mathbb{E}\left(M_1M_2\left[K(0)-N_1\right]\left[\int K\left(\frac{X_2-v}{h_{1,n}}\right)(f(v)-f(X_2))dv\right]\right). \tag{2.3.40}$$

Under the condition that $f(x)$ has a continuous bounded second order derivative, symmetric property of K and by taylor expansion and change of variable $u = \frac{X_2-v}{h_{1,n}}$, we have,

$$\int K\left(\frac{X_2-v}{h_{1,n}}\right)(f(v)-f(X_2))dv$$
$$= h_{1,n}\int K(u)\left(f(X_2-uh_{1,n})-f(X_2)\right)du = O(h_{1,n}^3).$$

Thus, by the boundedness of K and N_1 and (2.3.27),

$$(2.3.38) = O(h_{1,n}n^{-1}) = (2.3.40).$$

(2.3.39) $= 0$ since

$$\mathbb{E}\left(K\left(\frac{X_1-X_3}{h_{1,n}}\right)|X_1\right) = \int K\left(\frac{X_2-u}{h_{1,n}}\right)f(u)du = N_1.$$

Therefore, $Q_2 = O(h_{1,n}n^{-1})$. Notice that the first quantity in (2.3.29) is same as Q_3. Hence,

$$(2.3.18) - (2.3.23) = V_3(nh_{2,n})^{-1} + O(n^{-1}).$$

The difference between (2.3.20) and (2.3.25)

Next we denote $A(x) = K_t(x)\alpha(f(x))$ and $L_{t,2}(x) = L_{t,1}(x)\frac{\alpha'^2(f(x))}{\alpha(f(x))}$. Then, (2.1.6) and the decomposition of $\delta(t)$ in (2.1.7) give

$$
\begin{aligned}
(2.3.20) =& 2h_{2,n}^{-2}\mathbb{E}\left[A(X_1)L_{t,2}(X_2)(D_1(X_2;h_{1,n})+b_1(X_2;h_{1,n}))^2\right] + O(U^3(h_{1,n})) \\
=& 2h_{2,n}^{-2}\mathbb{E}\left[A(X_1)L_{t,2}(X_2)D_1^2(X_2;h_{1,n})\right] && (2.3.41)\\
&+4h_{2,n}^{-2}\mathbb{E}\left[A(X_1)L_{t,2}(X_2)D_1(X_2;h_{1,n})b_1(X_2;h_{1,n})\right] && (2.3.42)\\
&+2h_{2,n}^{-2}\mathbb{E}\left[A(X_1)L_{t,2}(X_2)b_1^2(X_2;h_{1,n})\right] + O(U^3(h_{1,n})) && (2.3.43)
\end{aligned}
$$

$$
\begin{aligned}
(2.3.41) =& \frac{2}{n^2h_{1,n}^2h_{2,n}^2}\mathbb{E}\left[A(X_1)L_{t,2}(X_2)\sum_{i=1}^{n}\left[K\left(\frac{X_2-X_i}{h_{1,n}}\right)-N_2\right]^2\right] && (2.3.44)\\
&+\frac{4}{n^2h_{1,n}^2h_{2,n}^2}\mathbb{E}\left[A(X_1)L_{t,2}(X_2)\sum_{1\le i<j\le n}\left[K\left(\frac{X_2-X_i}{h_{1,n}}\right)-N_2\right]\left[K\left(\frac{X_2-X_j}{h_{1,n}}\right)-N_2\right]\right] && (2.3.45)
\end{aligned}
$$

By independence of $X_i's$ and a similar argument as in (2.3.27), we have

$$
\begin{aligned}
(2.3.44) =& \frac{2}{n^2h_{1,n}^2h_{2,n}^2}\mathbb{E}\left[A(X_1)L_{t,2}(X_2)\left[K(0)-N_2\right]^2\right]\\
&+\frac{2}{n^2h_{1,n}^2h_{2,n}^2}\mathbb{E}\left[A(X_1)L_{t,2}(X_2)\left[K\left(\frac{X_2-X_1}{h_{1,n}}\right)-N_2\right]^2\right]\\
&+\frac{2(n-2)}{n^2h_{1,n}^2h_{2,n}^2}\mathbb{E}\left[A(X_1)L_{t,2}(X_2)\left[K\left(\frac{X_2-X_3}{h_{1,n}}\right)-N_2\right]^2\right] && (2.3.46)\\
=& O(nh_{1,n})^{-2} + (2.3.46)
\end{aligned}
$$

and

$$
\begin{aligned}
(2.3.45) = & \frac{4}{n^2 h_{1,n}^2 h_{2,n}^2} \mathbb{E}\left[A(X_1) L_{t,2}(X_2)\left[K\left(\frac{X_2 - X_1}{h_{1,n}}\right) - N_2\right][K(0) - N_2]\right] \\
& + \frac{4(n-2)}{n^2 h_{1,n}^2 h_{2,n}^2} \mathbb{E}\left[A(X_1) L_{t,2}(X_2)\left[K\left(\frac{X_2 - X_1}{h_{1,n}}\right) - N_2\right]\left[K\left(\frac{X_2 - X_3}{h_{1,n}}\right) - N_2\right]\right] \quad (2.3.47) \\
& + \frac{4(n-2)}{n^2 h_{1,n}^2 h_{2,n}^2} \mathbb{E}\left[A(X_1) L_{t,2}(X_2)[K(0) - N_2]\left[K\left(\frac{X_2 - X_3}{h_{1,n}}\right) - N_2\right]\right] \quad (2.3.48) \\
& = O(nh_{1,n})^{-2}.
\end{aligned}
$$

The last two terms (2.3.47) and (2.3.48) equal to 0 by arguments similar to (2.3.37). Now, we decompose the term (2.3.25). Again, by (2.1.6) and the decomposition of $\delta(t)$ in (2.1.7),

$$
\begin{aligned}
(2.3.25) = & 2h_{2,n}^{-2} \mathbb{E}[A(X_1)] \mathbb{E}\left[L_{t,2}(X_2)(D_1(X_2; h_{1,n}) + b_1(X_2; h_{1,n}))^2\right] + O(U^3(h_{1,n})) \\
= & 2h_{2,n}^{-2} \mathbb{E}[A(X_1)] \mathbb{E}\left[L_{t,2}(X_2) D_1^2(X_2; h_{1,n})\right] \quad (2.3.49) \\
& + 4h_{2,n}^{-2} \mathbb{E}[A(X_1)] \mathbb{E}\left[L_{t,2}(X_2)(D_1(X_2; h_{1,n}) b_1(X_2; h_{1,n}))\right] \quad (2.3.50) \\
& + 2h_{2,n}^{-2} \mathbb{E}[A(X_1)] \mathbb{E}\left[L_{t,2}(X_2) b_1^2(X_2; h_{1,n})\right] + O(U^3(h_{1,n})). \quad (2.3.51)
\end{aligned}
$$

$$
(2.3.49) = \frac{2}{n^2 h_{1,n}^2 h_{2,n}^2} \mathbb{E}A(X_1) \mathbb{E}\left[L_{t,2}(X_2) \sum_{i=1}^{n}\left[K\left(\frac{X_2 - X_i}{h_{1,n}}\right) - N_2\right]^2\right] \quad (2.3.52)
$$

$$
+ \frac{4}{n^2 h_{1,n}^2 h_{2,n}^2} \mathbb{E}A(X_1) \mathbb{E}\left[L_{t,2}(X_2) \sum_{1 \le i < j \le n}\left[K\left(\frac{X_2 - X_i}{h_{1,n}}\right) - N_2\right]\left[K\left(\frac{X_2 - X_j}{h_{1,n}}\right) - N_2\right]\right] \quad (2.3.53)
$$

By a similar argument as in (2.3.27), we have

$$
(2.3.52) = \frac{2}{n^2 h_{1,n}^2 h_{2,n}^2} \mathbb{E}[A(X_1)] \mathbb{E}\left[L_{t,2}(X_2)[K(0) - N_2]^2\right]
$$

$$
\begin{aligned}
&+\frac{2}{n^2h_{1,n}^2h_{2,n}^2}\mathbb{E}[A(X_1)]\mathbb{E}\left[L_{t,2}(X_2)\left[K\left(\frac{X_2-X_1}{h_{1,n}}\right)-N_2\right]^2\right]\\
&+\frac{2(n-2)}{n^2h_{1,n}^2h_{2,n}^2}\mathbb{E}[A(X_1)]\mathbb{E}\left[L_{t,2}(X_2)\left[K\left(\frac{X_2-X_3}{h_{1,n}}\right)-N_2\right]^2\right] \qquad (2.3.54)\\
=&O(n^{-2}h_{1,n}^{-2})+(2.3.54)
\end{aligned}
$$

and

$$
\begin{aligned}
(2.3.53)=&\frac{4}{n^2h_{1,n}^2h_{2,n}^2}\mathbb{E}[A(X_1)]\mathbb{E}\left[L_{t,2}(X_2)\left[K\left(\frac{X_2-X_1}{h_{1,n}}\right)-N_2\right][K(0)-N_2]\right]\\
&+\frac{4(n-2)}{n^2h_{1,n}^2h_{2,n}^2}\mathbb{E}[A(X_1)]\mathbb{E}\left[L_{t,2}(X_2)\left[K\left(\frac{X_2-X_1}{h_{1,n}}\right)-N_2\right]\right.\\
&\quad\left.\left[K\left(\frac{X_2-X_3}{h_{1,n}}\right)-N_2\right]\right] \qquad (2.3.55)\\
&+\frac{4(n-2)}{n^2h_{1,n}^2h_{2,n}^2}\mathbb{E}[A(X_1)]\mathbb{E}\left[L_{t,2}(X_2)[K(0)-N_2]\left[K\left(\frac{X_2-X_3}{h_{1,n}}\right)-N_2\right]\right] \qquad (2.3.56)\\
=&O(n^{-2}h_{1,n}^{-2}).
\end{aligned}
$$

The last two terms (2.3.55) and (2.3.56) are equal to 0. By similar explanation as Q_2, (2.3.50) $=O(h_{1,n}n^{-1})=$ (2.3.42). Also notice that (2.3.43) $=$ (2.3.51) and (2.3.46) $=$ (2.3.54). Therefore, we get,

$$
(2.3.20)-(2.3.25)=o\left(n^{-1}\right).
$$

The difference between (2.3.21) and (2.3.26)

By Proposition 2 of [10], for the ideal estimator (1.2.1), we have

$$
||\bar{f}(t;h_{2,n})-\mathbb{E}(\bar{f}(t;h_{2,n}))||_\infty=\sqrt{\frac{\log n}{nh_{2,n}}}.
$$

Therefore,

$$
||K_t(X_2)\alpha(f(X_2))-\mathbb{E}\left(K_t(X_2)\alpha(X_2)\right)||_\infty=h_{2,n}\sqrt{\frac{\log n}{nh_{2,n}}}.
$$

Hence, we have the following.

$$\begin{aligned}
(2.3.21) - (2.3.26) =& 2h_{2,n}^{-2}\mathbb{E}\left[K_t(X_2)\alpha(f(X_2))\alpha(f(X_1))\delta_2(t, X_1)\right] \\
& -2h_{2,n}^{-2}\mathbb{E}\left[K_t(X_2)\alpha(f(X_2))\right]\mathbb{E}\left[\alpha(f(X_1))\delta_2(t, X_1)\right] \\
=& 2h_{2,n}^{-2}\mathbb{E}\{\mathbb{E}\left[K_t(X_2)\alpha(f(X_2))\alpha(f(X_1))\delta_2(t, X_1)|X_2\right]\} \\
& -2h_{2,n}^{-2}\mathbb{E}\left[K_t(X_2)\alpha(f(X_2))\right]\mathbb{E}\{\mathbb{E}\left[\alpha(f(X_1))\delta_2(t, X_1)|X_2\right]\} \\
=& 2h_{2,n}^{-2}\mathbb{E}\Big[\Big(K_t(X_2)\alpha(f(X_2)) \\
& \quad - \mathbb{E}\left[K_t(X_2)\alpha(f(X_2))\right]\Big)\mathbb{E}\left[\alpha(f(X_1))\delta_2(t, X_1)|X_2\right]\Big] \\
\leq & \frac{2}{h_{2,n}}\sqrt{\frac{\log n}{nh_{2,n}}}\sup_{t,x\in\mathbb{R}}|\delta_2(t,x)| \\
=& \sqrt{\frac{\log n}{nh_{2,n}^3}}\; O\left(\left(h_{1,n}^2 + \sqrt{\frac{\log n}{nh_{1,n}}}\right)^2\right) = o(n^{-1}).
\end{aligned}$$

The difference between (2.3.19) and (2.3.24)

We shall apply the decomposition of $\delta(t)$ as in (2.1.7). The difference between (2.3.19) and (2.3.24) with the second part of (2.1.7) has order $o(n^{-1})$ by similar arguments as in the difference between (2.3.21) and (2.3.26). Therefore,

$$\begin{aligned}
(2.3.19) - (2.3.24) =& 2h_{2,n}^{-2}\mathbb{E}\left[K_t(X_2)\alpha(f(X_2))M_1(D_1(X_1;h_{1,n}) + b_1(X_1;h_{1,n}))\right] \\
& -2h_{2,n}^{-2}\mathbb{E}\left[K_t(X_2)\alpha(X_2)\right]\mathbb{E}\left[M_1(D_1(X_1;h_{1,n}) + b_1(X_1;h_{1,n}))\right] + o(n^{-1}) \\
=& 2h_{2,n}^{-2}\mathbb{E}\left[K_t(X_2)\alpha(f(X_2))M_1D_1(X_1;h_{1,n})\right] \\
& -2h_{2,n}^{-2}\mathbb{E}\left[K_t(X_2)\alpha(X_2)\right]\mathbb{E}\left[M_1D_1(X_1;h_{1,n})\right] + o(n^{-1}) \\
=& \frac{2}{nh_{1,n}h_{2,n}^2}\mathbb{E}\left[K_t(X_2)\alpha(f(X_2))M_1\left(K\left(\frac{X_1 - X_2}{h_{1,n}}\right) - N_1\right)\right] \qquad (2.3.57) \\
& -\frac{2}{nh_{1,n}h_{2,n}^2}\mathbb{E}\left[K_t(X_2)\alpha(X_2)\right]\mathbb{E}\left[M_1\left(K\left(\frac{X_1 - X_2}{h_{1,n}}\right) - N_1\right)\right] + o(n^{-1}) \qquad (2.3.58)
\end{aligned}$$

$$
\begin{aligned}
=&\frac{2}{nh_{1,n}h_{2,n}^2}\mathbb{E}\left[K_t(X_2)\alpha(f(X_2))M_1K\left(\frac{X_1-X_2}{h_{1,n}}\right)\right] && (2.3.59)\\
-&\frac{2}{nh_{1,n}h_{2,n}^2}\mathbb{E}\left[K_t(X_2)\alpha(f(X_2))\right]\mathbb{E}(M_1N_1)+o(n^{-1}). && (2.3.60)
\end{aligned}
$$

We have the equality (2.3.57) since X_i's are identical. The term (2.3.58) is zero.

Since K, L, α' and f are bounded functions and $K(v)$ has bounded support, we have

$$
\begin{aligned}
|V_2| &= \frac{2}{h_{1,n}h_{2,n}}\left|\mathbb{E}\left[K_t(X_2)\alpha(f(X_2))\int L_t(u)\alpha'(f(u))K\left(\frac{u-X_2}{h_{1,n}}\right)f(u)du\right]\right|\\
&= \frac{2}{h_{2,n}}\left|\mathbb{E}\left[K_t(X_2)\alpha(f(X_2))\int L_t(X_2+h_{1,n}v)\alpha'(f(X_2+h_{1,n}v))K(v)f(X_2+h_{1,n}v)dv\right]\right|\\
&\le \frac{C}{h_{2,n}}\mathbb{E}\left[K_t(X_2)\alpha(f(X_2))\right]=O\left(1\right) \qquad (2.3.61)
\end{aligned}
$$

and (2.3.59) $= O(nh_{2,n})^{-1}$. On the other hand,

$$
\mathbb{E}\left[K_t(X_2)\alpha(f(X_2))\right]=O(h_{2,n})
$$

and by change of variable and boundedness of L, α, f, and compact support of L,

$$
\begin{aligned}
\mathbb{E}(M_1N_1) =& \int L_t(x)\alpha'(f(x))f(x)\int K\left(\frac{x-v}{h_{1,n}}\right)f(v)dvdx\\
=&h_{1,n}\int L_t(x)\alpha'(f(x))f(x)\int K(w)f(x-h_{1,n}w)dwdx\\
=&h_{1,n}h_{2,n}\int L\left(u\alpha(f(t-uh_{2,n}))\right)\alpha'(f(t-uh_{2,n}))f(t-uh_{2,n})\\
&\times\int K(w)f(t-uh_{2,n}-h_{1,n}w)dwdu=O(h_{1,n}h_{2,n}).
\end{aligned}
$$

Therefore, in (2.3.60),

$$
\frac{2}{nh_{1,n}h_{2,n}^2}\mathbb{E}\left[K_t(X_2)\alpha(f(X_2))\right]\mathbb{E}(M_1N_1)=O(n^{-1}). \qquad (2.3.62)
$$

Hence, (2.3.61) and (2.3.62) give us

$$(2.3.19) - (2.3.24) = V_2(nh_{2,n})^{-1} + O(n^{-1}).$$

Thus, for $d = 1$,

$$Var(\hat{f}(t; h_{1,n}, h_{2,n})) = (V_1 + V_2 + V_3)(nh_{2,n})^{-1}(1 + o(1)).$$

$\square$

The asymptotic variance of the true estimator (1.2.8) for multidimensional case, d, is in Section 2.4.

2.4 CENTRAL LIMIT THEOREM FOR VARIABLE BANDWIDTH KERNEL DENSITY ESTIMATOR

2.4 Ideal Estimator

This subsection proposes the central limit theorem for the ideal estimator, $\bar{f}(t; h_{2,n})$. First, we state the following well known theorems.

Theorem 2.4.1 (Slutsky's Theorem [14]). *Let $\{X_n\}$, $\{A_n\}$, and $\{B_n\}$ be sequence of random variables. If $X_n \xrightarrow{D} X$, $A_n \xrightarrow{p} a$, and $B_n \xrightarrow{p} b$, where X is a random variable, and a and b are constants, then*

$$A_n + B_n X_n \xrightarrow{D} a + bX.$$

Theorem 2.4.2 (Lindeberg's Central Limit Theorem [3]). *Suppose that for each n, $X_{n,1}$, $\cdots$, X_{n,r_n} are independent. Put $S_n = X_{n,1} + \cdots + X_{n,r_n}$. Suppose that $\mathbb{E}(X_{n,k}) = 0$, $\sigma^2_{n,k} = \mathbb{E}(X^2_{n,k})$, and $s^2_n = \sum_{k=1}^{r_n} \sigma^2_{n,k}$. Then, under the condition*

$$\lim_{n\to\infty} \sum_{k=1}^{r_n} \frac{1}{s_n^2} \int_{|X_{n,k}|\geq \epsilon s_n} X^2_{n,k} dP = 0$$

for $\epsilon > 0$, we have

$$\frac{S_n}{s_n} \xrightarrow{D} N(0,1).$$

Proposition 2.4.3. *Let $\alpha(f(t)) = cp^{1/2}(c^{-2}f(t))$ for some $c > 0$, and p be as in Assumptions 2. Define $\bar{f}(t;h_n)$ by equation (1.2.1). Assume $h_n = c_2 n^{-1/(8+d)}$ for some $c_2 > 0$. Then, under Assumptions 2,*

$$\sqrt{nh_{2,n}^d}[\bar{f}(t;h_n) - f(t)] \xrightarrow{D} N\left(c_2^{(d+8)/2} \sum_{|v|=4} \tau_v D_v(1/f)/v!, \alpha^d(f(t))f(t)\mu_0\right).$$

Proof. The ideal estimator $\bar{f}(t;h_n)$ in (1.2.1) can be written as a sample mean of triangular array for random variables, i.e., $\bar{f}(t;h_n) = \bar{Y} = \frac{1}{n}\sum_{i=1}^n Y_{n,i}$, where

$$Y_{n,i} = \frac{1}{h_n^d} K(h_n^{-1}\alpha(f(X_i))(t - X_i))\alpha^d(f(X_i)). \tag{2.4.1}$$

By (2.3.7) and Proposition 2.3.2,

$$\sqrt{Var(\bar{f}(t;h_n))} = (1 + o(1))\sqrt{\frac{1}{nh_n^d}\alpha^d(f(t))f(t)\mu_0}.$$

Hence, by Theorem 2.4.2 for triangular array of random variables,

$$\sqrt{nh_n^d}[\bar{f}(t;h_n) - \mathbb{E}\bar{f}(t;h_n)] \xrightarrow{D} N(0, \alpha^d(f(t))f(t)\mu_0).$$

Since $\mathbb{E}\bar{f}(t;h_n) - f(t) = \left(\sum_{|v|=4} \tau_v D_v(1/f)/v!\right) h_n^4(1 + o(1))$ by [16, 17] or Corollary 1 in [10], if we take $h_n = c_2 n^{-1/(8+d)}$ for some constant $c_2 > 0$,

$$\sqrt{nh_n^d}[\mathbb{E}\bar{f}(t;h_n) - f(t)] = c_2^{(d+8)/2} \sum_{|v|=4} \tau_v D_v(1/f)/v!(1 + o(1)),$$

Note that,

$$\bar{f}(t;h_n) - f(t) = \bar{f}(t;h_n) - \mathbb{E}\bar{f}(t;h_n) + \mathbb{E}\bar{f}(t;h_n) - f(t).$$

Thus, by Theorem 2.4.1, for $t \in D_r$,

$$\sqrt{nh_n^d}[\bar{f}(t;h_n) - f(t)] \xrightarrow{D} N\left(c_2^{(d+8)/2}\sum_{|v|=4}\tau_v D_v(1/f)/v!, \alpha^d(f(t))f(t)\mu_0\right).$$

$\square$

2.4 True Estimator

Based on the above central limit theorem for the ideal estimator, we have the following.

Theorem 2.4.4. *Let $X_1, ..., X_n$ be a random sample of size n from a population with density function $f(t)$, $t \in \mathbb{R}^d$. Consider $\hat{f}(t;h_{1,n},h_{2,n})$ defined in* (1.2.8)*. The function $\alpha(x)$ in the estimator $\hat{f}(t;h_{1,n},h_{2,n})$ is defined in (1.2.2) for a nondecreasing clipping function $p(s)$. Let $h_{2,n} = c_2 n^{-1/(8+d)}$ for some constants $c_2 > 0$ and assume that $U(h_{1,n}) = o(h_{2,n}^2)$. Then, for $t \in \mathcal{D}_r$, under Assumptions 1 for K, and Assumptions 2 for p and f,*

$$\sqrt{nh_{2,n}^d}[\hat{f}(t;h_{1,n},h_{2,n}) - \mathbb{E}\hat{f}(t;h_{1,n},h_{2,n})] \xrightarrow{D} N\left(0, \sigma_t^2\right) \tag{2.4.2}$$

and

$$\sqrt{nh_{2,n}^d}[\hat{f}(t;h_{1,n},h_{2,n}) - f(t)] \xrightarrow{D} N\left(c_2^{(d+8)/2}\sum_{|v|=4}\tau_v D_v(1/f)/v!, \sigma_t^2\right), \tag{2.4.3}$$

where $\sigma_t^2 = \alpha^d(f(t))f(t)\mu_0 + f^3(t)\frac{[(\alpha^d)'(f(t))]^2}{d^2\alpha^d(f(t))}\int_{\mathbb{R}^d} L^2(z)dz + f^2(t)(\alpha^d)'(f(t))\mu_0$, $\mu_0 = \int_{\mathbb{R}^d} K^2(u)du$, and $L(x) = K(x) + xK'(x)$.

The proof of the Theorem 2.4.4 needs the following lemma.

Lemma 2.4.5. *Under the Assumptions from theorem 2.4.4, we have that*

$$\mathbb{E}R_{n,1} = \mathbb{E}\bar{f}(t; h_{2,n}) = f(t) + h_{2,n}^4 \sum_{|v|=4} \tau_v D_v(1/f)/v! + o(h_{2,n}^4),$$

$$h_{2,n}^d Var(R_{n,1}) \xrightarrow{n\to\infty} \sigma_t^2, \tag{2.4.4}$$

where $R_{n,1} = Y_{n,1} + Z_{n,1}$, $Z_{n,1} = \frac{1}{dh_{1,n}^d h_{2,n}^d}[\mathbb{E}_X H_t(X, X_1) - \mathbb{E}H_t]$, $Y_{n,1}$ *is defined as (2.4.1), and* $H_t(\cdot,\cdot)$ *as (2.2.10).*

Proof. By corollary 2.2.1, we have that

$$\mathbb{E}\bar{f}(t; h_{2,n}) = f(t) + h_{2,n}^4 \sum_{|v|=4} \tau_v D_v(1/f)/v! + o(h_{2,n}^4).$$

Also note that $\mathbb{E}Z_{n,1} = 0$. Thus,

$$\mathbb{E}R_{n,1} = \mathbb{E}\bar{f}(t; h_{2,n}) = f(t) + h_{2,n}^4 \sum_{|v|=4} \tau_v D_v(1/f)/v! + o(h_{2,n}^4).$$

Now we will show (2.4.4). By the definition of variance,

$$\begin{aligned} h_{2,n}^d Var(R_{n,1}) &= h_{2,n}^d[\mathbb{E}R_{n,1}^2 - (\mathbb{E}R_{n,1})^2] \\ &= h_{2,n}^d \mathbb{E}Y_{n,1}^2 + h_{2,n}^d \mathbb{E}Z_{n,1}^2 + 2h_{2,n}^d \mathbb{E}(Y_{n,1}Z_{n,1}) - h_{2,n}^d(\mathbb{E}R_{n,1})^2. \end{aligned}$$

Since $\frac{1}{h_{2,n}^d} A(X_1) = Y_{n,1}$, by (2.3.8), we have that,

$$h_{2,n}^d \mathbb{E}Y_{n,1}^2 = \alpha^d(f(t))f(t)\mu_0 + o(1).$$

We only need to calculate the limit of terms $h_{2,n}^d \mathbb{E}(Y_{n,1}Z_{n,1})$ and $h_{2,n}^d \mathbb{E}Z_{n,1}^2$ to show (2.4.4).

Let $x_1 = x - uh_{1,n}$ and $x_1 = t - vh_{2,n}$. Then,

$$\begin{aligned}
&\frac{1}{h_{1,n}^{2d} h_{2,n}^{d}} \mathbb{E}(\mathbb{E}_X H_t(X, X_1))^2 \\
&= \frac{1}{h_{1,n}^{2d} h_{2,n}^{d}} \int_{\mathbb{R}^d} \left[\int_{\mathbb{R}^d} L\left(\frac{t-x}{h_{2,n}} \alpha(f(x)) \right) (\alpha^d)'(f(x)) K\left(\frac{x - x_1}{h_{1,n}} \right) f(x) dx \right]^2 f(x_1) dx_1 \\
&= \frac{1}{h_{2,n}^{d}} \int_{\mathbb{R}^d} \Big[\int_{\mathbb{R}^d} L\left(\frac{t - x_1 - h_{1,n} u}{h_{2,n}} \alpha(f(x_1 + h_{1,n} u)) \right) (\alpha^d)'(f(x_1 + h_{1,n} u)) K(u) \\
&\times f(x_1 + h_{1,n} u) du \Big]^2 f(x_1) dx_1 \\
&= \int_{\mathbb{R}^d} \Big[\int_{\mathbb{R}^d} L\left(\left(v - \frac{h_{1,n} u}{h_{2,n}} \right) \alpha(f(t - h_{2,n} v + h_{1,n} u)) \right) (\alpha^d)'(f(t - h_{2,n} v + h_{1,n} u)) \\
&\times K(u) f(t - h_{2,n} v + h_{1,n} u) du \Big]^2 f(t - h_{2,n} v) dv.
\end{aligned}$$

Now, by the boundedness of K, α, L_1, f, the compact support of K, and since $h_{2,n} = c_2 n^{-1/(8+d)}$ and $U(h_{1,n}) = o(h_{2,n}^2)$, we have

$$\frac{1}{h_{1,n}^{2d} h_{2,n}^{d}} \mathbb{E}(\mathbb{E}_X H_t(X, X_1))^2 \xrightarrow{n \to \infty} f^3(t) \frac{[(\alpha^d)'(f(t))]^2}{\alpha^d(f(t))} \int_{\mathbb{R}^d} L^2(z) dz.$$

Also,

$$\begin{aligned}
&\frac{1}{h_{1,n}^{d}} \mathbb{E}[Y_{n,1} \mathbb{E}_X (H_t(X, X_1))] \\
&= \frac{1}{h_{1,n}^{d} h_{2,n}^{d}} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} L\left(\frac{t-x}{h_{2,n}} \alpha(f(x)) \right) (\alpha^d)'(f(x)) K\left(\frac{x - x_1}{h_{1,n}} \right) K\left(\frac{t - x_1}{h_{2,n}} \alpha(f(x_1)) \right) \\
&\times (\alpha^d)(f(x_1)) f(x) f(x_1) dx dx_1 \\
&= \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} L\left(\left(v - \frac{u h_{1,n}}{h_{2,n}} \right) \alpha(f(u h_{1,n} + t - v h_{2,n})) \right) (\alpha^d)'(f(u h_{1,n} + t - v h_{2,n})) K(u) \\
&\times K(v \alpha(f(t - h_{2,n} v))) (\alpha^d)(f(t - h_{2,n} v)) f(t - h_{2,n} v) f(u h_{1,n} + t - v h_{2,n}) du dv.
\end{aligned}$$

By integration by parts and compact support of K, we have $\int_{\mathbb{R}^d} K(t) \sum_{i=1}^d t_i K_i'(t) dt = -d\mu_0/2$. Also, by the boundedness of K, α, L_1, f, the compact support of K, and since

$h_{2,n} = c_2 n^{-1/(8+d)}$ and $U(h_{1,n}) = o(h_{2,n}^2)$, we have

$$\frac{1}{h_{1,n}^d}\mathbb{E}[Y_{n,1}\mathbb{E}_X(H_t(X,X_1))] \xrightarrow{n\to\infty} \frac{1}{2}df^2(t)(\alpha^d)'(f(t))\mu_0. \tag{2.4.5}$$

By the change of variables $y = x - uh_{1,n}+$ and $y = t - vh_{2,n}$, we have that

$$\begin{aligned}\mathbb{E}(H_t) &= \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} L\left(\frac{t-x}{h_{2,n}}\alpha(f(x))\right)(\alpha^d)'(f(x))K\left(\frac{x-y}{h_{1,n}}\right)f(x)f(y)dxdy\\ &= h_{1,n}^d h_{2,n}^d \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} L\left(\left(v - \frac{uh_{1,n}}{h_{2,n}}\right)\alpha(f(uh_{1,n}+t-vh_{2,n}))\right)\\ &\times (\alpha^d)'(f(uh_{1,n}+t-vh_{2,n}))K(u)f(t-h_{2,n}v)f(uh_{1,n}+t-vh_{2,n})dudv.\end{aligned}$$

By the boundedness of L, α, f, and compact support of L, we have that $\mathbb{E}(H_t)$ is bounded by $Ch_{1,n}^d h_{2,n}^d$ for some $C > 0$. Thus, $\frac{1}{h_{1,n}^{2d}h_{2,n}^d}[\mathbb{E}(H_t)]^2 \to 0$ as $n \to \infty$. Hence,

$$\begin{aligned}h_{2,n}^d\mathbb{E}Z_{n,1}^2 &= \frac{1}{d^2h_{1,n}^{2d}h_{2,n}^d}\mathbb{E}(\mathbb{E}_X H_t(X,X_1))^2 - \frac{1}{d^2h_{1,n}^{2d}h_{2,n}^d}[\mathbb{E}(H_t)]^2\\ &\xrightarrow{n\to\infty} f^3(t)\frac{[(\alpha^d)'(f(t))]^2}{d^2\alpha^d(f(t))}\int_{\mathbb{R}^d} L^2(z)dz\end{aligned} \tag{2.4.6}$$

and

$$\begin{aligned}h_{2,n}^d\mathbb{E}(Y_{n,1}Z_{n,1}) &= \frac{1}{dh_{1,n}^d}\mathbb{E}[Y_{n,1}\mathbb{E}_X(H_t(X,X_1))] - \frac{1}{dh_{1,n}^d}\mathbb{E}Y_{n,1}\mathbb{E}H_t\\ &\xrightarrow{n\to\infty} \frac{1}{2}f^2(t)(\alpha^d)'(f(t))\mu_0.\end{aligned} \tag{2.4.7}$$

Thus,

$$h_{2,n}^d Var(R_{n,1}) \xrightarrow{n\to\infty} \sigma_t^2.$$

$\square$

Proof of Theorem 2.4.4. $\hat{f}(t; h_{1,n}, h_{2,n})$ in (1.2.8) has decomposition

$$\hat{f}(t; h_{1,n}, h_{2,n}) - \mathbb{E}\hat{f}(t; h_{1,n}, h_{2,n}) = \hat{f}(t; h_{1,n}, h_{2,n}) - \bar{f}(t; h_{2,n}) \tag{2.4.8}$$

$$+ \bar{f}(t; h_{2,n}) - \mathbb{E}\bar{f}(t; h_{2,n}) \tag{2.4.9}$$

$$+ \mathbb{E}\bar{f}(t; h_{2,n}) - \mathbb{E}\hat{f}(t; h_{1,n}, h_{2,n}), \tag{2.4.10}$$

$$\begin{aligned} \hat{f}(t; h_{1,n}, h_{2,n}) - f(t) &= \hat{f}(t; h_{1,n}, h_{2,n}) - \mathbb{E}\hat{f}(t; h_{1,n}, h_{2,n}) \\ &+ \mathbb{E}\hat{f}(t; h_{1,n}, h_{2,n}) - f(t). \end{aligned} \tag{2.4.11}$$

Since

$$\sqrt{nh_{2,n}^d}[\mathbb{E}\bar{f}(t; h_{2,n}) - \mathbb{E}\hat{f}(t; h_{1,n}h_{2,n})] = o(h_{2,n}^4\sqrt{nh_{2,n}^d}) = o(1) \tag{2.4.12}$$

by the analysis in Section 2.2.3 and when $h_{2,n} = c_2 n^{-1/(8+d)}$, the term (2.4.10) is negligible in the central limit theorems (2.4.2) and (2.4.3).

The term (2.4.8) has decomposition as in (2.1.13). By Lemma 2.2.3, we have that all the terms in the decomposition (2.1.13) has order $o_{a.s.}(h_{2,n}^4)$ or small except for the terms $\mathbb{E}(\bar{f}; h_{2,n})$ and $\frac{1}{nh_{2,n}^d}\sum_{i=1}^n L\left(\frac{t-X_i}{h_{2,n}}\alpha(f(X_i))\right)\alpha^d(f(X_i))\delta(X_i)$. By similar explanation as that of (2.4.12) and Theorem 2.4.1, they are negligible in the proof of the Theorem 2.4.4 when multiplied by $\sqrt{nh_{2,n}^d}$.

We can further decompose the term $\frac{1}{nh_{2,n}^d}\sum_{i=1}^n L\left(\frac{t-X_i}{h_{2,n}}\alpha(f(X_i))\right)\alpha^d(f(X_i))\delta(X_i)$ into the random variation part D_1 and the bias b_1 by the decomposition (2.1.7):

$$\begin{aligned} &\frac{1}{nh_{2,n}^d}\sum_{i=1}^n L\left(\frac{t-X_i}{h_{2,n}}\alpha(f(X_i))\right)\alpha^d(f(X_i))\delta(X_i) \\ &= \frac{1}{dnh_{2,n}^d}\sum_{i=1}^n \left[L\left(\frac{t-X_i}{h_{2,n}}\alpha(f(X_i))\right)(\alpha^d)'(f(X_i))D(X_i; h_{1,n})\right] \end{aligned} \tag{2.4.13}$$

$$+\frac{1}{dnh_{2,n}^d}\sum_{i=1}^n\left[L\left(\frac{t-X_i}{h_{2,n}}\alpha(f(X_i))\right)(\alpha^d)'(f(X_i))b(X_i;h_{1,n})\right] \tag{2.4.14}$$

$$+\frac{1}{2nh_{2,n}^d}\sum_{i=1}^n\left[L\left(\frac{t-X_i}{h_{2,n}}\alpha(f(X_i))\right)\alpha^{d-1}(f(X_i))(\alpha^d)''(\eta(X_i))[\hat{f}(X_i;h_{1,n})-f(X_i)]^2\right]. \tag{2.4.15}$$

By (3.24) of [10], we have that (2.4.15) $= O_{a.s.}(U^2(h_{1,n})) = o_{a.s.}(h_{2,n}^4)$ and by the analysis for (3.27) of the same paper, we have (2.4.14) $= o_{a.s.}(h_{2,n}^4)$. Also, the term (2.4.13) multiplied by d can be further decomposed into (2.2.12), (2.2.13) and (2.2.14) from Section 2.2.3, and by the order of (2.2.13) given by (3.33) of [10], (2.2.13) $= o_{a.s.}(h_{2,n}^4)$. Next we show that (2.2.12) $= o_p(h_{2,n}^4)$.

We have that $\mathbb{E}[H_t(X_1,X_1)-\mathbb{E}_Y H_t(X_1,Y)]$ and $\mathbb{E}[H_t(X_1,X_1)-\mathbb{E}_Y H_t(X_1,Y)]^2$ are bounded by $Ch_{2,n}^d$, for some constant C with some change of variables. Thus,

$$\begin{aligned}
&\mathbb{E}(2.2.12)^2 := \mathbb{E}B_n^2\\
&= \frac{1}{n^3h_{1,n}^{2d}h_{2,n}^{2d}}\mathbb{E}[H_t(X_1,X_1)-\mathbb{E}_Y H_t(X_1,Y)]^2 + \frac{n-1}{n^3h_{1,n}^{2d}h_{2,n}^{2d}}[\mathbb{E}(H_t(X_1,X_1)-\mathbb{E}_Y H_t(X_1,Y))]^2\\
&\le \frac{C}{n^2h_{1,n}^{2d}}, \text{for some constant C.}
\end{aligned}$$

Let $\epsilon > 0$. By Markov's inequality when $h_{2,n} = c_2 n^{-1/(8+d)}$ and $U(h_{1,n}) = o(h_{2,n}^2)$, we have

$$\mathbb{P}\left(\left|\frac{B_n}{h_{2,n}^4}\right| > \epsilon\right) \le \frac{\mathbb{E}(B_n^2)}{\epsilon^2 h_{2,n}^8} \le \frac{C}{n^2h_{1,n}^{2d}\epsilon^2h_{2,n}^8} \xrightarrow{n\to\infty} 0.$$

Hence, (2.2.12)$=o_p(h_{2,n}^4)$.

By the above analysis, only the term $\bar{R}-\mathbb{E}\bar{f}(t;h_{2,n})$ have contribution in the central limit theorems (2.4.2) and (2.4.3), where $\bar{R} = \frac{1}{n}\sum_{i=1}^n R_{n,i}$. The other terms are of order $o_{a.s.}(h_{2,n}^4)$. Thus, they are negligible by similar explanation as (2.4.12). To prove the central limit theorem (2.4.2), it suffices to derive a central limit theorem for $\bar{R} - \mathbb{E}\bar{f}(t; h_{2,n})$ where $\bar{R}$ is the sample mean of triangular array of random variables $R_{n,i}$, $1 \leq i \leq n$.

By Theorem 2.4.2 and Lemma 2.4.5, we have

$$\sqrt{nh_{2,n}^d}[\bar{R} - \mathbb{E}\bar{R}] \xrightarrow{D} N\left(0, \sigma_t^2\right)$$

and by Theorem 2.4.1,

$$\sqrt{nh_{2,n}^d}[\hat{f}(t; h_{1,n}, h_{2,n}) - \mathbb{E}\hat{f}(t; h_{1,n}, h_{2,n})] \xrightarrow{D} N\left(0, \sigma_t^2\right).$$

Since the term (2.4.11) $= \mathbb{E}\hat{f}(t; h_{1,n}, h_{2,n}) - f(t) = c_2^{(d+8)/2} \sum_{|v|=4} \tau_v D_v(1/f)/v! h_{2,n}^4 (1 + o(1))$ by Proposition 2.2.2,

$$\sqrt{nh_{2,n}^d}[\mathbb{E}\hat{f}(t; h_{1,n}, h_{2,n}) - f(t)] = c_2^{(d+8)/2} \sum_{|v|=4} \tau_v D_v(1/f)/v!(1 + o(1)).$$

Thus, for $t \in \mathcal{D}_r$ when $h_{2,n} = c_2 n^{-1/(8+d)}$ and $U(h_{1,n}) = o(h_{2,n}^2)$, and under Assumptions 1 for K, and Assumptions 2 for p and f,

$$\sqrt{nh_{2,n}^d}[\hat{f}(t; h_{1,n}, h_{2,n}) - f(t)] \xrightarrow{D} N\left(c_2^{(d+8)/2} \sum_{|v|=4} \tau_v D_v(1/f)/v!, \sigma_t^2\right).$$

$\square$

2.5 OPTIMAL BANDWIDTH AND SIMULATION ON VARIABLE BANDWIDTH KERNEL DENSITY ESTIMATOR

2.5 Optimal Bandwidth $h^*_{2,n}$ on Variable Bandwidth Kernel Density Estimation

In this section we evaluate the performance of the variable bandwidth kernel density estimator (VKDE), (1.2.8), in one dimensional case. Instead of the true estimator (1.2.8), Jones *et al.* [15] did simulation study for the ideal estimator (1.2.1) in one dimensional case. First of all, we provide a result on the integrated mean squared error (IMSE) of the VKDE and therefore a formula of optimal bandwidth.

Theorem 2.5.1. *Under the conditions in Proposition 2.2.2 and Theorem 2.4.4, the IMSE on D_r is*

$$R(h_{1,n}, h_{2,n})|D_r = h^8_{2,n} \int_{D_r} (\sum_{|v|=4} \tau_v D_v(1/f)/v!)^2 dt + \frac{1}{nh^d_{2,n}} \int_{D_r} \sigma^2_t dt + o(h^8_{2,n}), \qquad (2.5.1)$$

Furthermore, the optimal bandwidth $h^*_{2,n}$ is given by

$$h^*_{2,n} = \left[\frac{n \int_{D_r} (\sum_{|v|=4} \tau_v D_v(1/f)/v!)^2 dt}{\int_{D_r} \sigma^2_t dt} \right]^{-1/(8+d)}, \qquad (2.5.2)$$

where $\sigma^2_t = \alpha^d(f(t))f(t)\mu_0 + f^3(t)\frac{[(\alpha^d)'(f(t))]^2}{d^2\alpha^d(f(t))} \int_{\mathbb{R}^d} L^2(z)dz + f^2(t)(\alpha^d)'(f(t))\mu_0$, $\mu_0 = \int_{\mathbb{R}^d} K^2(u)du$, and $L(x) = K(x) + xK'(x)$.

Proof. From the analysis of Theorem 2.4.4 and Proposition 2.2.2, it is clear that $Var(\hat{f}(t; h_{1,n}, h_{2,n})) = \frac{(1+o(1))\sigma^2_t}{nh^d_{2,n}}$ and $(bias)^2 = \sum_{|v|=4} \tau_v D_v(1/f)/v!)^2$ for $t \in D_r$. Thus, we have (2.5.1). □

2.5 Simulation on Variable Bandwidth Kernel Density Estimation

We compare the performance of VKDE and KDE by conducting simulation study of t-distribution ($t_4(0,1)$), Cauchy(0,1) and Pareto(0,1).

The sample size is $n = 50,000$ for each simulation study. For all the simulations, we use KDE as in (1.1.1) with the normal kernel function. We use the code *density()* in the programming software R and the default bandwidth chosen by R in the estimation for $t_4(0,1)$. For Cauchy(0,1) or Pareto(0,1), the code *density()* in R can not provide a classical kernel density estimate. Instead, we make new code and select the bandwidth which optimizes the performance among a variety of bandwidths. For VKDE, we assume that $h_{1,n} = n^{-1/5}$, $h_{2,n} = n^{-1/9}$, and use the Tricube kernel:

$$K(u) = \frac{70}{81}(1 - |u|^3)^3 1_{|u|\le 1}$$

in either the pilot kernel density estimator or the true estimator (1.2.8). The following five time differentiable clipping function p with $t_0 = 2$ ([10]) is applied:

$$p(t) = \begin{cases} 1 + \frac{t^6}{64}\left(1 - 2(t-2) + \frac{9}{4}(t-2)^2 - \frac{7}{4}(t-2)^3 + \frac{7}{8}(t-2)^4\right) & \text{if } 0 \le t \le 2 \\ t & \text{if } t \ge 2 \\ 1 & \text{if } t \le 0 \end{cases}.$$

The simulation study in Figure 2.5.1 shows that, for each of these three distributions, VKDE has better performance than KDE, especially in the tail area.

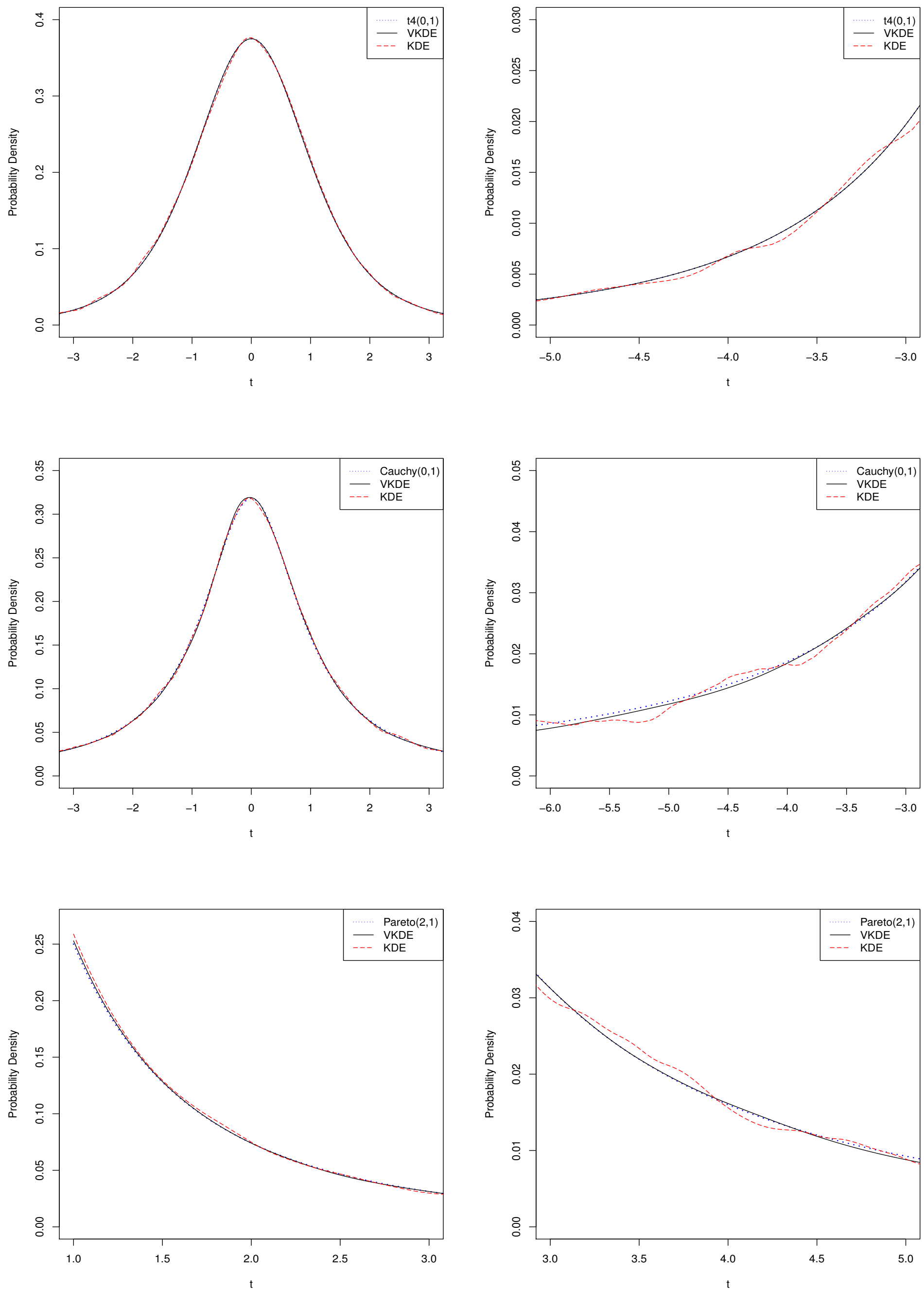

Figure 2.5.1: The probability density functions of t-distribution ($t_4(0,1)$), Cauchy(0,1) and Pareto(0,1), the kernel density estimates (KDE), and the variable kernel density estimates (VKDE) with 50,000 observations generated from t-distribution ($t_4(0,1)$), Cauchy(0,1) and Pareto(0,1) distribution. The left one shows the estimate in the main area with the mode. The right one shows the estimate in the tail area.

3 VARIABLE BANDWIDTH KERNEL REGRESSION ESTIMATION

Chapter 3 is devoted to developments of variable bandwidth kernel regression estimators.

3.1 DECOMPOSITION FOR KERNEL REGRESSION ESTIMATION

Let $X_i \in \mathbb{R}$.

Similar to δ in Section 2.1, define $\delta_r(t) = \delta_r(t, n)$ by the equation

$$\delta_r(t) = \frac{\alpha(\hat{g}(t; h_{1,n})) - \alpha(g(t))}{\alpha(g(t))}.$$

Then,

$$\alpha(\hat{g}(t; h_{1,n})) = \alpha(g(t))(1 + \delta_r(t)) \tag{3.1.1}$$

Moreover, we have

$$\delta_r(t) = \frac{\alpha'(g(t))[\hat{g}(t; h_{1,n}) - g(t)]}{\alpha(g(t))} + \frac{\alpha''(\gamma)[\hat{g}(t; h_{1,n}) - g(t)]^2}{2\alpha(g(t))} \tag{3.1.2}$$

where $\gamma = \gamma(t) \geq 0$ is between $\hat{g}(t; h_{1,n})$ and $g(t)$. Note that, since $p \geq 1$ and if p' and p'' are uniformly bounded on $[0, \infty)$, we have $|\alpha''(\gamma(t, h_{1,n}))| \leq c^{-3} A_2$ for some constant A_2 that does not depend on n or t but only on p.

Define $D_2(t;h_{1,n}) = \hat{g}(t;h_{1,n}) - \mathbb{E}(\hat{g}(t;h_{1,n}))$ and $b_2(t;h_{1,n}) = \mathbb{E}(\hat{g}(t;h_{1,n})) - g(t)$. By [24] (Theorem 3.3) and [5],

$$\sup_{t\in\mathbb{R}} b_2(t;h_{1,n}) = O(h_{1,n}^2), \tag{3.1.3}$$

$$\text{and } \sup_{t\in\mathbb{R}} |D_2(t;h_{1,n})| = O_{a.s.}\left(\sqrt{\frac{\log n}{nh_{1,n}}}\right). \tag{3.1.4}$$

Note that for $d = 1$, by (2.1.5), $\sqrt{\frac{\log n}{nh_{1,n}}} + h_{1,n}^2 = U(h_{1,n})$. Then, we have,

$$\sup_{t\in\mathbb{R}} |\delta_r(t)| = \sup_{t\in\mathbb{R}} |\hat{g}(t;h_{1,n}) - g(t)| = \sup_{t\in\mathbb{R}} |D_2(t;h_{1,n}) + b_2(t;h_{1,n})| = O_{a.s.}\left(U(h_{1,n})\right), \tag{3.1.5}$$

for $f, r \in \mathcal{P}_{C,2}$.

Recall

$$L_1(t) = tK'(t) \quad \text{and} \quad L(t) = K(t) + L_1(t), \quad t \in \mathbb{R}. \tag{3.1.6}$$

We then have the following Taylor expansion

$$\begin{aligned} K\left(\frac{t - X_i}{h_{2,n}}\alpha(\hat{g}(X_i;h_{1,n}))\right) &= K\left(\frac{t - X_i}{h_{2,n}}\alpha(g(X_i))\right) \\ + K'\left(\frac{t - X_i}{h_{2,n}}\alpha(g(X_i))\right) &\frac{(t - X_i)}{h_{2,n}}\alpha(g(X_i))\delta_r(X_i) + \delta_3(t;X_i), \end{aligned} \tag{3.1.7}$$

where

$$\delta_3(t, X_i) = K''(\xi)\frac{(t - X_i)^2}{2h_{2,n}^2}\alpha^2(g(X_i))\delta_r^2(X_i),$$

ξ being a number between $\frac{t-X_i}{h_{2,n}}\alpha(g(X_i))$ and $\frac{t-X_i}{h_{2,n}}\alpha(g(X_i))(1+\delta_r(X_i))$. By the similar analysis as that for δ_2 in Section 2.1

$$\sup_{t,x\in\mathbb{R}} |\delta_3(t,x)| = O_{\text{a.s.}}\left(\|\hat{g}(\cdot;h_{1,n}) - g(\cdot)\|_\infty^2\right) = O_{a.s.}(U^2(h_{1,n})) \tag{3.1.8}$$

if $f, r \in \mathcal{P}_{C,2}$. Therefore, by using the expansion (3.1.1) of $\alpha(\hat{f})$, the Taylor expansion (3.1.7), and the notations L_1 and L in (3.1.6), we get

$$
\begin{aligned}
\hat{g}(t; h_{1,n}, h_{2,n}) =& \bar{g}(t; h_{2,n}) \\
&+ \frac{1}{nh_{2,n}} \sum_{i=1}^{n} L\left(\frac{t - X_i}{h_{2,n}} \alpha(g(X_i))\right) \alpha(g(X_i)) \delta_r(X_i) Y_i \\
&+ \frac{1}{nh_{2,n}} \sum_{i=1}^{n} L_1\left(\frac{t - X_i}{h_{2,n}} \alpha(g(X_i))\right) \alpha(g(X_i)) \delta_r^2(X_i) Y_i \\
&+ \frac{1}{nh_{2,n}} \sum_{i=1}^{n} \alpha(g(X_i)) \delta_3(t, X_i) \delta_r(X_i) Y_i \\
&+ \frac{1}{nh_{2,n}} \sum_{i=1}^{n} \alpha(g(X_i)) \delta_3(t, X_i) Y_i. \qquad (3.1.9)
\end{aligned}
$$

3.2 BIAS OF VARIABLE BANDWIDTH KERNEL REGRESSION ESTIMATOR

3.2 Ideal Estimator

For $r_1 > 0$, $1 \le t_0 < \infty$, and $c > 0$, define $D_{rf} = D_1 \cap D_2$, where

$$
\begin{aligned}
D_1 &= \{t \in \mathbb{R} : f(t) > r_1 > t_0 c^2,\ |t| < 1/r_1\}, \\
D_2 &= \{t \in \mathbb{R} : g(t) > r_1 > t_0 c^2,\ |t| < 1/r_1\}
\end{aligned}
$$

Recall that

$$
\bar{r}(t; h_n) \quad = \quad \frac{\frac{1}{nh_n} \sum_{i=1}^{n} K\left(\frac{t - X_i}{h_n} \alpha(g(X_i))\right) \alpha(g(X_i)) Y_i}{\frac{1}{nh_n} \sum_{i=1}^{n} K\left(\frac{t - X_i}{h_n} \alpha(f(X_i))\right) \alpha(f(X_i))} = \frac{\bar{g}(t; h_n)}{\bar{f}(t; h_n)}. \qquad (3.2.1)
$$

Consider bias of the ideal variable bandwidth regression estimator (3.2.1) at a fixed point $t \in D_{rf}$. The rate of convergence of this bias depends on the properties of f, r, and the clipping function p.

We give the following assumptions needed for the theorem and lemmas to follow.

Assumptions 3. *The sequence h_n satisfies the following classical conditions:*

$$h_n \searrow 0, \quad \frac{nh_n}{|\log h_n|} \to \infty, \quad \frac{|\log h_n|}{\log\log n} \to \infty, \quad \text{and } nh_n \nearrow \infty,$$

as $n \to \infty$. K is non-negative, symmetric about zero, and has bounded and continuous second order derivatives. Assume that K has support $[-T, T]$ for some $T < \infty$. Set $\alpha(x) = cp^{1/2}(c^{-2}x)$ for $x \in \mathbb{R}$ and some $c > 0$. Moreover, let r be bounded function with $r \in C^4(\mathbb{R})$.

The following proposition and its proof are contained in [17] and [15] (see also the theorem in [11]). We require this proposition to prove Theorem 3.2.5, more specifically to work on $\mathbb{E}(\bar{g})$ and $\mathbb{E}(\bar{f})$.

Proposition 3.2.1 ([17]). *Assume that K is a symmetric kernel and has bounded support in $[-T, T]$. Let η be a function in $C^l(\mathbb{R})$ and ξ a function in $C^{l+1}(\mathbb{R})$. Assume $\xi(t) \geq c > 0$ for some $c > 0$ and all $t \in \mathbb{R}$. Then,*

$$\frac{1}{h}\int K\left(\frac{t-s}{h}\xi(s)\right)\xi(s)\eta(s)ds = \sum_{k=0}^{l} a_k(t)h^k + o(h^l)$$

as $h \to 0$, uniformly for $t \in \mathbb{R}$, and the set of functions $a_k(t)$, which are uniformly bounded and equicontinuous, are defined as

$$a_{2k+1}(t) = 0, \quad a_{2k}(t) = \frac{\rho_{2k}}{(2k)!}D_{2k}\left(\frac{\eta(t)}{\xi^{2k}(t)}\right),$$

for $k \leq l/2$, in particular, $a_0(t) = \eta(t)$. Here $\rho_k = \int w^k K(w)dw$ and D_k is the k^{th} derivative.

The following three lemmas are necessary to establish our results on the bias of $\bar{r}(t; h_n)$.

Lemma 3.2.2 ([10])**.** *Under Assumptions* 3*,*

$$\sup_{t\in\mathbb{R}} \left|\bar{f}(t;h_n) - \mathbb{E}\bar{f}(t;h_n)\right| = O_{a.s}\left(\sqrt{\frac{\log n}{nh_n}}\right), \tag{3.2.2}$$

$$\sup_{t\in\mathbb{R}} \left|\bar{g}(t;h_n) - \mathbb{E}\bar{g}(t;h_n)\right| = O_{a.s}\left(\sqrt{\frac{\log n}{nh_n}}\right). \tag{3.2.3}$$

(3.2.2) is given as Proposition 2 in [10]. The proof of the relation (3.2.3) is similar to that of (3.2.2) because Y_1 is bounded.

Lemma 3.2.3. *Under Assumptions* 3*, for $t \in D_{rf}$,*

$$\frac{\mathbb{E}(\bar{g}(t;h_n))}{\mathbb{E}(\bar{f}(t;h_n))} = r(t) + \frac{\rho_4 h_n^4}{24}\left[\frac{1}{f(t)}D_4\left(\frac{1}{g(t)}\right) - \frac{g(t)}{f^2(t)}D_4\left(\frac{1}{f(t)}\right)\right] + o(h_n^4)$$

as $n \to \infty$.

Proof. By the definition of $\bar{g}(t;h_n)$,

$$\begin{aligned}\mathbb{E}(\bar{g}(t;h_n)) &= \frac{1}{h_n}\int\int K\left(\frac{t-x}{h_n}\alpha(g(x))\right)\alpha(g(x))yf(y|x)f(x)dydx \\ &= \frac{1}{h_n}\int K\left(\frac{t-x}{h_n}\alpha(g(x))\right)\alpha(g(x))r(x)f(x)dx.\end{aligned}$$

By Proposition 3.2.1 using $\alpha(g(x)) = \xi(x)$ and $\eta(x) = r(x)f(x)$, we have

$$\begin{aligned}\mathbb{E}(\bar{g}(t;h_n)) &= r(t)f(t) + \frac{\rho_2 h_n^2}{2}\frac{d^2}{dt^2}\left(\frac{r(t)f(t)}{\alpha^2(g(t))}\right) \\ &+ \frac{\rho_4}{24}D_4\left(\frac{f(t)r(t)}{\alpha^4(g(t))}\right)h_n^4 + o(h_n^4).\end{aligned} \tag{3.2.4}$$

For $t \in D_{rf}$, $\alpha(g(t)) = cp^{1/2}(f(t)r(t)/c^2) = f^{1/2}(t)r^{1/2}(t)$ by Proposition 3.2.1, $a_2(t) = 0$ on D_{rf}. Thus, by Corollary 2.2.1 and Proposition 3.2.1, we have

$$\begin{aligned}\frac{\mathbb{E}(\bar{g}(t;h_n))}{\mathbb{E}(\bar{f}(t;h_n))} &= \frac{f(t)r(t) + \frac{\rho_4}{24}D_4\left(\frac{1}{f(t)r(t)}\right)h_n^4 + o(h_n^4)}{f(t) + \frac{\rho_4}{24}D_4\left(\frac{1}{f(t)}\right)h_n^4 + o(h_n^4)} \\ &= r(t) + \frac{\rho_4 h_n^4}{24}\left[\frac{1}{f(t)}D_4\left(\frac{1}{g(t)}\right) - \frac{g(t)}{f^2(t)}D_4\left(\frac{1}{f(t)}\right)\right] + o(h_n^4).\end{aligned} \tag{3.2.5}$$

The last equality uses the first order Taylor expansion at 0 of $\frac{1}{1+\frac{\rho_4}{24f(t)}D_4\left(\frac{1}{f(t)}\right)h_n^4+o(h_n^4)}$. This concludes the proof. □

Lemma 3.2.4. *Let $\frac{\log n}{nh_n} = o(h_n^4)$. Under Assumptions 3, for $t \in D_{rf}$,*

$$\mathbb{E}(\bar{r}(t;h_n)) = \mathbb{E}\left(\frac{\bar{g}(t;h_n)}{\bar{f}(t;h_n)}\right) = \frac{\mathbb{E}(\bar{g}(t;h_n))}{\mathbb{E}(\bar{f}(t;h_n))} + o(h_n^4) \text{ as } n \to \infty.$$

Proof. Using the formula

$$\frac{1}{z} = 1 - (z-1) + \cdots + (-1)^p(z-1)^p + (-1)^{p+1}\frac{(z-1)^{p+1}}{z}. \tag{3.2.6}$$

with $p = 2$ and $z = \bar{f}/\mathbb{E}(\bar{f})$, we get

$$\begin{aligned}\mathbb{E}(\bar{r}(t;h_n)) = \mathbb{E}\left(\frac{\bar{g}(t;h_n)}{\bar{f}(t;h_n)}\right) &= \mathbb{E}\left(\frac{\bar{g}(t;h_n)}{\frac{\bar{f}(t;h_n)}{\mathbb{E}(\bar{f}(t;h_n))}\mathbb{E}(\bar{f}(t;h_n))}\right) \\ &= \frac{\mathbb{E}(\bar{g}(t;h_n))}{\mathbb{E}(\bar{f}(t;h_n))} \\ &\quad - \frac{Cov(\bar{g}(t;h_n), \bar{f}(t;h_n))}{(\mathbb{E}(\bar{f}(t;h_n)))^2} && (3.2.7) \\ &\quad + \frac{\mathbb{E}\Big(\bar{g}(t;h_n)\Big(\bar{f}(t;h_n) - \mathbb{E}(\bar{f}(t;h_n))\Big)^2\Big)}{(\mathbb{E}(\bar{f}(t;h_n))^3} && (3.2.8) \\ &\quad - \frac{\mathbb{E}\Big(\bar{r}(t;h_n)\Big(\bar{f}(t;h_n) - \mathbb{E}(\bar{f}(t;h_n))\Big)^3\Big)}{(\mathbb{E}(\bar{f}(t;h_n))^3}. && (3.2.9)\end{aligned}$$

First we will handle the numerator of the term (3.2.7). We will show the following identity.

$$Cov(\bar{f}(t;h_n),\bar{g}(t;h_n)) = \frac{1}{nh_n}\beta_n(t) + O(n^{-1}), \tag{3.2.10}$$

$$\text{where } \beta_n(t) := \int K\left(u\alpha(g(t-h_nu))\right)\alpha(g(t-h_nu))$$
$$K\left(u\alpha(f(t-h_nu))\right)\alpha(f(t-h_nu))g(t-h_nu)du.$$

Using the fact that (X_i, Y_i) are *i.i.d.*, we have

$$\begin{aligned}
&Cov(\bar{f}(t;h_n),\bar{g}(t;h_n))\\
&= \frac{1}{n^2h_n^2}Cov\left(\sum_{i=1}^{n} K\left(\frac{t-X_i}{h_n}\alpha(g(X_i))\right)\alpha(g(X_i))Y_i, \sum_{i=1}^{n} K\left(\frac{t-X_i}{h_n}\alpha(f(X_i))\right)\alpha(f(X_i))\right)\\
&= \frac{1}{nh_n^2}Cov\left(K\left(\frac{t-X_1}{h_n}\alpha(g(X_1))\right)\alpha(g(X_1))Y_1, K\left(\frac{t-X_1}{h_n}\alpha(f(X_1))\right)\alpha(f(X_1))\right).
\end{aligned} \tag{3.2.11}$$

By definition $Cov(W_1,W_2) = \mathbb{E}(W_1W_2) - \mathbb{E}W_1\mathbb{E}W_2$. So, we will show $\mathbb{E}(W_1W_2)$ part of (3.2.11) using integration by substitution with $u = (t-s)/h_n$.

$$\begin{aligned}
&\mathbb{E}\left(K\left(\frac{t-X_1}{h_n}\alpha(g(X_1))\right)\alpha(g(X_1))Y_1 K\left(\frac{t-X_1}{h_n}\alpha(f(X_1))\right)\alpha(f(X_1))\right)\\
&= \int\int K\left(\frac{t-s}{h_n}\alpha(g(s))\right)\alpha(g(s))vK\left(\frac{t-s}{h_n}\alpha(f(s))\right)\alpha(f(s))f(s,v)dvds\\
&= \int K\left(\frac{t-s}{h_n}\alpha(g(s))\right)\alpha(g(s))K\left(\frac{t-s}{h_n}\alpha(f(s))\right)\alpha(f(s))r(s)f(s)ds\\
&= \int K\left(\frac{t-s}{h_n}\alpha(g(s))\right)\alpha(g(s))K\left(\frac{t-s}{h_n}\alpha(f(s))\right)\alpha(f(s))g(s)ds\\
&= h_n\int K\left(u\alpha(g(t-h_nu))\right)\alpha(g(t-h_nu))K\left(u\alpha(f(t-h_nu))\right)\alpha(f(t-h_nu))g(t-h_nu)du\\
&= h_n\beta_n(t).
\end{aligned} \tag{3.2.12}$$

$\mathbb{E}W_1\mathbb{E}W_2$ part of (3.2.11) can be derived from the Propostion 3.2.1. Using Proposition 3.2.1 with $\xi(t) = \alpha(g(t))$ and $\eta(t) = g(t)$, we have

$$\mathbb{E}\left(K\left(\frac{t-X_1}{h_n}\alpha(g(X_1))\right)\alpha(g(X_1))Y_1\right) = h_n\mathbb{E}(\bar{g}(t;h_n)) = h_n(g(t) + O(h_n^2)). \tag{3.2.13}$$

Again, using the Proposition 3.2.1 with $\xi(t) = \alpha(f(t))$ and $\eta(t) = f(t)$,

$$\mathbb{E}\left(K\left(\frac{t-X_1}{h_n}\alpha(f(X_1))\right)\alpha(f(X_1))\right) = h_n\mathbb{E}(\bar{f}(t;h_n)) = h_n(f(t) + O(h_n^2)). \tag{3.2.14}$$

Hence, by (3.2.11), (3.2.12), (3.2.13), and (3.2.14), we have (3.2.10). Also, $\beta_n(t)$ is bounded since K, α, g, and f are bounded. Thus, (3.2.7) $= O((nh_n)^{-1})$.

Now, we will show that terms (3.2.8) and (3.2.9) are of order $o(h_n^4)$. For $t \in D_{rf}$ by Corollary 2.2.1 and Lemma 3.2.2, we have that $\frac{1}{\bar{f}(t;h_n)} \xrightarrow{a.s.} \frac{1}{f(t)}$. Similarly, by (3.2.5) and Lemma 3.2.2 we have $\bar{g}(t;h_n) \xrightarrow{a.s.} g(t)$. Since $\bar{r} = \bar{g}/\bar{f}$, we have $\bar{r}(t;h_n) \xrightarrow{a.s.} r(t)$. Therefore, by (3.2.5), Corollary 2.2.1 and Lemma 3.2.2,

$$(3.2.8), (3.2.9) = o(h_n^4).$$

This concludes the proof of Lemma 3.2.4. $\square$

The idea of the proof of the following theorem is to treat the numerator $\bar{g}$ and the denominator $\bar{f}$ separately and go through the fraction using the formula (3.2.6). This idea is similar to the proof of the bias of the Nadaraya-Watson Estimator in [24].

Theorem 3.2.5. *Let (X_i, Y_i), $1 \le i \le n$, be i.i.d. random vectors. Assume that Y_1 is bounded and $\frac{\log n}{nh_n} = o(h_n^4)$. Then, under Assumptions 2 and 3, for $t \in D_{rf}$,*

$$\mathbb{E}(\bar{r}(t;h_n)) = r(t) + \frac{\rho_4 h_n^4}{24}\left[\frac{1}{f(t)}D_4\left(\frac{1}{g(t)}\right) - \frac{g(t)}{f^2(t)}D_4\left(\frac{1}{f(t)}\right)\right] + o(h_n^4) \tag{3.2.15}$$

as $n \to \infty$, with $\rho_4 = \int w^4 K(w)dw$, $g(t) = f(t)r(t)$, and D_4 is the fourth order derivative.

Proof. By Lemma 3.2.4, we get

$$\mathbb{E}(\bar{r}(t;h_n)) = \mathbb{E}\left(\frac{\bar{g}(t;h_n)}{\bar{f}(t;h_n)}\right) = \frac{\mathbb{E}(\bar{g}(t;h_n))}{\mathbb{E}(\bar{f}(t;h_n))} + o(h_n^4)$$

and the theorem follows from Lemma 3.2.3. $\square$

3.2 True Estimator

Recall that the true estimator of $r(t)$ is:

$$\begin{aligned}
\hat{r}(t;h_{1,n},h_{2,n}) &= \frac{\frac{1}{nh_{2,n}}\sum_{i=1}^{n} K\left(\frac{t-X_i}{h_{2,n}}\alpha(\hat{g}(X_i;h_{1,n}))\right)\alpha(\hat{g}(X_i;h_{1,n}))Y_i}{\frac{1}{nh_{2,n}}\sum_{i=1}^{n} K\left(\frac{t-X_i}{h_{2,n}}\alpha(\hat{f}(X_i;h_{1,n}))\right)\alpha(\hat{f}(X_i;h_{1,n}))} \\
&= \frac{\hat{g}(t;h_{1,n},h_{2,n})}{\hat{f}(t;h_{1,n},h_{2,n})}. \qquad (3.2.16)
\end{aligned}$$

We need to approximate the bias for the true variable bandwidth kernel regression estimator. We first investigate the numerator of $\hat{r}(t;h_{1,n},h_{2,n})$.

Proposition 3.2.6. *Define $\hat{g}(t;h_{1,n},h_{2,n})$ by (1.4.5). Assume that $U(h_{1,n}) = o(h_{2,n}^2)$ and $\frac{\log n}{nh_{2,n}} = o(h_{2,n}^4)$. Then, under Assumptions 3 with $h_n = h_{2,n}$ and Assumptions 2, for $t \in \mathcal{D}_{rf}$,*

$$\mathbb{E}(\hat{g}(t;h_{1,n},h_{2,n})) = g(t) + \frac{\rho_4 h_{2,n}^4}{24} D_4\left(\frac{1}{g(t)}\right) + o(h_{2,n}^4) \text{ as } n \to \infty.$$

Proof. Consider the bounds on (3.1.9). By analogous argument for Lemma 2.2.3 or (3.20) of [10], (3.1.5), (3.1.8), and the boundedness of α, K, L_1, and Y_1, for all $t \in D_{rf}$,

$$\begin{aligned}
&\sup_{t\in\mathbb{R}}\left|\frac{1}{nh_{2,n}}\sum_{i=1}^{n}\left[L_1\left(\frac{t-X_i}{h_{2,n}}\alpha(g(X_i))\right)\alpha(g(X_i))\delta_r^2(X_i)Y_i\right.\right. \\
&\left.\left.+\,\alpha(g(X_i))\delta_3(t,X_i)\delta_r(X_i)Y_i + \alpha(g(X_i))\delta_3(t,X_i)Y_i\right]\right| \qquad (3.2.17) \\
&= O_{a.s}(U^2(h_{1,n})) = o_{a.s.}(h_{2,n}^4).
\end{aligned}$$

Hence, by (3.1.9) and (3.2.17),

$$\begin{aligned}\mathbb{E}(\hat{g}(t;h_{1,n},h_{2,n})) =& \mathbb{E}(\bar{g}(t;h_{2,n})) + o(h_{2,n}^4)\\ &+\frac{1}{h_{2,n}}\mathbb{E}\left(L\left(\frac{t-X_1}{h_{2,n}}\alpha(g(X_1))\right)\alpha(g(X_1))\delta_r(X_1)Y_1\right).\end{aligned} \tag{3.2.18}$$

Now reformulate (3.2.18) using the expansion (3.1.2) of δ_r, and the decomposition of $\hat{g}-g$ into variance D_2 and bias b_2.

$$\frac{1}{h_{2,n}}\mathbb{E}\left(L\left(\frac{t-X_1}{h_{2,n}}\alpha(g(X_1))\right)\alpha(g(X_1))\delta_r(X_1)Y_1\right)$$

$$= \frac{1}{h_{2,n}}\mathbb{E}\left(L\left(\frac{t-X_1}{h_{2,n}}\alpha(g(X_1))\right)\alpha'(g(X_1))Y_1D_2(X_1;h_{1,n})\right) \tag{3.2.19}$$

$$+ \frac{1}{h_{2,n}}\int L\left(\frac{t-s}{h_{2,n}}\alpha(g(s))\right)\alpha'(g(s))b_2(s;h_{1,n})g(s)ds \tag{3.2.20}$$

$$+ \frac{1}{2h_{2,n}}\mathbb{E}\left[L\left(\frac{t-X_1}{h_{2,n}}\alpha(g(X_1))\right)\right.$$

$$\left.\times\, \alpha''(\gamma(X_1))Y_1[\hat{g}(X_1;h_{1,n})-g(X_1)]^2\right]. \tag{3.2.21}$$

By (3.1.5), the order of the term (3.2.21) is

$$|(3.2.21)| = O(U^2(h_{1,n})) = o(h_{2,n}^4). \tag{3.2.22}$$

For the estimation of (3.2.20), we denote $h_{2,n} = h_2$ and $h_{1,n} = h_1$. By similar explanation as the proof of Proposition 2.3.2 with $\gamma = \alpha$ and $\zeta = g$, we have the change of variables

$$h_2 z = (t-s)\alpha(g(t-(t-s))), \text{ that is, } t-s = V_t(h_2 z) \tag{3.2.23}$$

in the following integral is valid for all h_2 small enough:

$$\begin{aligned}
&\frac{1}{h_2}\int L\left(\frac{t-s}{h_2}\alpha(g(s))\right)\alpha'(g(s))b_2(s;h_1)g(s)ds\\
&=\frac{1}{h_2}\int L\left(\frac{t-s}{h_2}\alpha(g(s)))\right)\alpha'(g(s))g(s)\int\frac{1}{h_1}K\left(\frac{s-u}{h_1}\right)(g(u)-g(s))duds\\
&=\int\Bigg(\alpha'(g(t-V_t(h_2z))g(t-V_t(h_2z))\left(\frac{\partial V_t}{\partial v}\right)_{v=h_2z}\\
&\times\int K(y)\big(g(t-V_t(h_2z)-yh_1)-g(t-V_t(h_2z))\big)dy\Bigg)L(z)dz\\
&\qquad:=\int F(h_2z)G(h_2z)L(z)dz.
\end{aligned}$$

Then, using $\int L(z)dz=\int zL(z)dz=0$ and Taylor expansion of $(FG)(h_2z)$ at 0, we obtain

$$(3.2.20)=\frac{h_2^2}{2}\int\big((F''G+2F'G'+FG'')(\theta(h_2z))z^2\big)\,L(z)dz,$$

where $0\le\theta(h_2z)\le h_2z$.

Since F, F',F'', and L are bounded with L having a bounded support, we need to show that $\|G\|_\infty$, $\|G'\|_\infty$, $\|G''\|_\infty$ are all $O(h_1^2)$ for $f,r\in\mathcal{P}_{C,4}$. Note that

$$G(h_2z)=\int K(y)\left(g(t-V_t(h_2z)-yh_1)-g(t-V_t(h_2z))\right)dy. \tag{3.2.24}$$

By (3.1.3), we have $\|G\|_\infty=O(h_1^2)$.

Let's show that $\|G'\|_\infty=O(h_1^2)$ for $f,r\in\mathcal{P}_{C,4}$.

Notice that

$$G'(h_2z)=\left(\frac{\partial V_t}{\partial v}\right)_{v=h_2z}\int K(y)\Big(g'(t-V_t(h_2z)-yh_1)-g'(t-V_t(h_2z))\Big)dy.$$

Using Taylor expansion of g' at $t - V_t(h_2 z)$, we obtain

$$\begin{aligned}
&\left(\frac{\partial V_t}{\partial v}\right)_{v=h_2 z} \int K(y)\Big(g'(t - V_t(h_2 z) - yh_1) - g'(t - V_t(h_2 z))\Big)dy \\
&= \left(\frac{\partial V_t}{\partial v}\right)_{v=h_2 z} \int K(y) \\
&\times \Big(g'(t - V_t(h_2 z)) - yh_1 g''(t - V_t(h_2 z)) - g'(t - V_t(h_2 z)) + O(h_1^2)\Big)dy \\
&= \left(\frac{\partial V_t}{\partial v}\right)_{v=h_2 z} h_1 g''(t - V_t(h_2 z)) \int yK(y)dy + O(h_1^2) = O(h_1^2).
\end{aligned}$$

The last equality uses the symmetry of K. The proof to show that $\|G''\|_\infty$ is $O(h_1^2)$ for $f, r \in \mathcal{P}_{C,4}$ is very similar to the proof of $\|G'\|_\infty$. Therefore, since $U(h_1) = o(h_2^2)$ we conclude

$$|(3.2.20)| = O(h_1^2 h_2^2) = o(h_2^4) \quad \text{for } f, r \in \mathcal{P}_{C,4}. \tag{3.2.25}$$

Estimation of (3.2.19) requires U-processes. Let H be an integrable function of two $i.i.d.$ vectors (X, Y), where the vectors (X_i, Y_i) are $i.i.d.$ copies of (X, Y). We can set

$$H_t^g\big((X_1, Y_1), (X_2, Y_2)\big) := L\left(\frac{t - X_1}{h_{2,n}}\alpha(g(X_1))\right)\alpha'(g(X_1))Y_1 K\left(\frac{X_1 - X_2}{h_{1,n}}\right)Y_2. \tag{3.2.26}$$

Then, by U-statistics and second order Hoeffding projection as in (2.2.8) and (2.2.9), (3.2.19) is the expectation of the following term which can be decomposed into a diagonal term and a U-statistic term, as follows:

$$\begin{aligned}
&\frac{1}{nh_{2,n}} \sum_{i=1}^n L\left(\frac{t - X_i}{h_{2,n}}\alpha(g(X_i))\right)\alpha'(g(X_i))Y_i D_2(X_i; h_{1,n}) && (3.2.27) \\
&= \frac{1}{n^2 h_{1,n} h_{2,n}} \sum_{i=1}^n \Big(H_t^g\big((X_i, Y_i), (X_i, Y_i)\big) - \mathbb{E}_{X,Y} H_t^g\big((X_i, Y_i), (X, Y)\big)\Big) && (3.2.28)
\end{aligned}$$

$$+ \frac{n-1}{nh_{1,n}h_{2,n}} U_n \left(\pi_2(H_t^g(\cdot,\cdot))\right) \tag{3.2.29}$$

$$+ \frac{n-1}{n^2 h_{1,n}h_{2,n}} \sum_{i=1}^{n} \Big(\mathbb{E}_{X,Y} H_t^g\big((X,Y),(X_i,Y_i)\big) - \mathbb{E}H_t^g\Big). \tag{3.2.30}$$

By analogous arguments as (2.2.15), we have

$$\mathbb{E}U_n \left(\pi_2(H_t^g(\cdot,\cdot))\right) = \mathbb{E}(\mathbb{E}_{X_1,Y_1} H_t^g\big((X_1,Y_1),(X_2,Y_2)\big) - \mathbb{E}H_t^g) = 0.$$

Now, set $\bar{Q}_i(t) = H_t^g\big((X_i,Y_i),(X_i,Y_i)\big) - \mathbb{E}_{X,Y} H_t^g\big((X_i,Y_i),(X,Y)\big)$. By similar calculation as (2.2.16), observe that,

$$\mathbb{E}|\bar{Q}_1(t)| \le B_2 h_{2,n},$$

for some finite constant B_2. Then,

$$\begin{aligned}
(3.2.19) &= \mathbb{E}((3.2.27)) \\
&= \mathbb{E}\left(\frac{1}{n^2 h_{1,n}h_{2,n}} \sum_{i=1}^{n} (H_t^g\big((X_i,Y_i),(X_i,Y_i)\big) - \mathbb{E}_{X,Y} H_t^g\big((X_i,Y_i),(X,Y)\big)\right) \\
&= O\left(\frac{1}{nh_{1,n}}\right).
\end{aligned} \tag{3.2.31}$$

By (3.2.22), (3.2.25), and (3.2.31), (3.2.18) $= o(h_{2,n}^4)$. Thus,

$$\mathbb{E}(\hat{g}(t;h_{1,n},h_{2,n})) = \mathbb{E}(\bar{g}(t;h_{2,n})) + o(h_{2,n}^4) \tag{3.2.32}$$

For $t \in D_{rf}$, by (3.2.5) we have that $\mathbb{E}(\bar{g}(t;h_{2,n})) = g(t) + \frac{\rho_4 h_{2,n}^4}{24 f(t)} D_4\left(\frac{1}{g(t)}\right) + o(h_{2,n}^4)$, Hence, since $U(h_{1,n}) = o(h_{2,n}^2)$ and $\frac{\log n}{nh_{2,n}} = o(h_{2,n}^4)$, by boundedness of K, g, and α, we have

$$\mathbb{E}(\hat{g}(t;h_{1,n},h_{2,n})) = g(t) + \frac{\rho_4 h_{2,n}^4}{24 f(t)} D_4\left(\frac{1}{g(t)}\right) + o(h_{2,n}^4).$$

□

We will use the following notations for the rest of the chapter:

$$K_g(X_i) = K\left(\frac{t-X_i}{h_{2,n}}\alpha(g(X_i))\right), K_f(X_i) = K\left(\frac{t-X_i}{h_{2,n}}\alpha(f(X_i))\right),$$
$$L_g(X_i) = L\left(\frac{t-X_i}{h_{2,n}}\alpha(g(X_i))\right), \text{ and } L_f(X_i) = L\left(\frac{t-X_i}{h_{2,n}}\alpha(f(X_i))\right).$$

Also note that in the one dimensional case, we have the following decomposition for $\hat{f}(t;h_{1,n},h_{2,n}))$ using (2.1.2) and the Taylor series expansion of $K\left(\frac{t-X_i}{h_{2,n}}\alpha(\hat{f}(X_i;h_{1,n}))\right)$:

$$\begin{aligned}
\hat{f}(t;h_{1,n},h_{2,n}) =& \bar{f}(t;h_{2,n}) \\
&+\frac{1}{nh_{2,n}}\sum_{i=1}^{n} L\left(\frac{t-X_i}{h_{2,n}}\alpha(f(X_i))\right)\alpha(f(X_i))\delta(X_i) \\
&+\frac{1}{nh_{2,n}}\sum_{i=1}^{n} L_1\left(\frac{t-X_i}{h_{2,n}}\alpha(f(X_i))\right)\alpha(f(X_i))\delta^2(X_i) \\
&+\frac{1}{nh_{2,n}}\sum_{i=1}^{n} \alpha(f(X_i))\delta_2(t,X_i)\delta(X_i) \\
&+\frac{1}{nh_{2,n}}\sum_{i=1}^{n} \alpha(f(X_i))\delta_2(t,X_i).
\end{aligned} \tag{3.2.33}$$

To work on the bias of $\hat{r}(t;h_{1,n},h_{2,n})$, we require Proposition 3.2.6 and the following lemma.

Lemma 3.2.7. *Assume $U(h_{1,n}) = o(h_{2,n}^2)$ and $\frac{\log n}{nh_{2,n}} = o(h_{2,n}^4)$. Under Assumptions 2 and Assumptions 3 with $h_n = h_{2,n}$, for $t \in D_{rf}$,*

$$\mathbb{E}(\hat{r}(t;h_{1,n},h_{2,n})) = \mathbb{E}\left(\frac{\hat{g}(t;h_{1,n},h_{2,n})}{\hat{f}(t;h_{1,n},h_{2,n})}\right) = \frac{\mathbb{E}(\hat{g}(t;h_{1,n},h_{2,n}))}{\mathbb{E}(\hat{f}(t;h_{1,n},h_{2,n}))} + o(h_{2,n}^4) \tag{3.2.34}$$

as $n \to \infty$.

Proof. By formula (3.2.6) for $p = 2$ with $z = \hat{f}/\mathbb{E}\hat{f}$ [equivalent to form obtained in (3.2.7) – (3.2.9) for $\mathbb{E}(\bar{r}(t; h_n))$], we have

$$\begin{aligned}\mathbb{E}(\hat{r}(t; h_{1,n}, h_{2,n})) &= \frac{\mathbb{E}(\hat{g}(t; h_{1,n}, h_{2,n}))}{\mathbb{E}(\hat{f}(t; h_{1,n}, h_{2,n}))} \\ &- \frac{Cov(\hat{g}(t; h_{1,n}, h_{2,n}), \hat{f}(t; h_{1,n}, h_{2,n}))}{(\mathbb{E}(\hat{f}(t; h_{1,n}, h_{2,n})))^2} \qquad (3.2.35)\\ &+ \frac{\mathbb{E}\Big(\hat{g}(t; h_{1,n}, h_{2,n})\big(\hat{f}(t; h_{1,n}, h_{2,n}) - \mathbb{E}(\hat{f}(t; h_{1,n}, h_{2,n}))\big)^2\Big)}{(\mathbb{E}(\hat{f}(t; h_{1,n}, h_{2,n})))^3} \qquad (3.2.36)\\ &- \frac{\mathbb{E}\Big(\hat{r}(t; h_{1,n}, h_{2,n})\big(\hat{f}(t; h_{1,n}, h_{2,n}) - \mathbb{E}(\hat{f}(t; h_{1,n}, h_{2,n}))\big)^3\Big)}{\big(\mathbb{E}(\hat{f}(t; h_{1,n}, h_{2,n}))\big)^3}. \qquad (3.2.37)\end{aligned}$$

We need to show that (3.2.35), (3.2.36), and (3.2.37) are of order $o(h_{2,n}^4)$. The steps to prove this is similar to proof of Lemma 3.2.4. First, we work on the numerator of the term (3.2.35). By the definition of covariance, we have

$$\begin{aligned}&Cov\Big(\hat{g}(t; h_{1,n}, h_{2,n}), \hat{f}(t; h_{1,n}, h_{2,n})\Big)\\ &= \mathbb{E}(\hat{g}(t; h_{1,n}, h_{2,n})\hat{f}(t; h_{1,n}, h_{2,n})) - \mathbb{E}(\hat{g}(t; h_{1,n}, h_{2,n}))\mathbb{E}(\hat{f}(t; h_{1,n}, h_{2,n})).\end{aligned}$$

We show that $\mathbb{E}(\hat{g}(t; h_{1,n}, h_{2,n})\hat{f}(t; h_{1,n}, h_{2,n})) = \mathbb{E}(\bar{g}(t; h_{2,n})\bar{f}(t; h_{2,n})) + o(h_{2,n}^4)$. By the decomposition (3.2.33) for $\hat{f}(t; h_{1,n}, h_{2,n})$ and decomposition (3.1.9) for $\hat{g}(t; h_{1,n}, h_{2,n})$, we get:

$$\begin{aligned}&\mathbb{E}(\hat{g}(t; h_{1,n}, h_{2,n})\hat{f}(t; h_{1,n}, h_{2,n}))\\ =&\frac{1}{n^2 h_{2,n}^2}\mathbb{E}\Bigg(\Bigg[\sum_{i=1}^n K_g(X_i)\alpha(g(X_i))Y_i + \sum_{i=1}^n L_g(X_i)\alpha(g(X_i))\delta_r(X_i)Y_i\\ &+ \sum_{i=1}^n L_1\left(\frac{t - X_i}{h_{2,n}}\alpha(g(X_i))\right)\alpha(g(X_i))\delta_r^2(X_i)Y_i\\ &+ \sum_{i=1}^n \alpha(g(X_i))\delta_3(t, X_i)\delta_r(X_i)Y_i + \sum_{i=1}^n \alpha(g(X_i))\delta_3(t, X_i)Y_i\Bigg]\end{aligned}$$

$$\left[\sum_{i=1}^{n} K_f(X_i)\alpha(f(X_i)) + \sum_{i=1}^{n} L_f(X_i)\alpha(f(X_i))\delta(X_i)\right.$$
$$+\sum_{i=1}^{n} L_1\left(\frac{t-X_i}{h_{2,n}}\alpha(f(X_i))\right)\alpha(f(X_i))\delta^2(X_i)$$
$$\left.\left.+\sum_{i=1}^{n}\alpha(f(X_i))\delta_2(t,X_i)\delta(X_i)+\sum_{i=1}^{n}\alpha(f(X_i))\delta_2(t,X_i)\right]\right).$$

By (2.1.6), (2.1.12), (3.1.5), and (3.1.8), notice that all the terms in $\mathbb{E}(\hat{g}(t;h_{1,n},h_{2,n})\hat{f}(t;h_{1,n},h_{2,n}))$ that have multiple of $\delta_r(\cdot)\delta(\cdot)$, $\delta^2(\cdot)$, $\delta_r^2(\cdot)$, $\delta_2(\cdot)$, or $\delta_3(\cdot)$ have order $o(h_{2,n}^4)$ or smaller if we take $U(h_{1,n}) = o(h_{2,n}^2)$. Thus, by the boundedness of Y_1, K, L_1, L, f, r and α:

$$\mathbb{E}(\hat{g}(t;h_{1,n},h_{2,n})\hat{f}(t;h_{1,n},h_{2,n})) = \mathbb{E}(\bar{g}(t;h_{2,n})\bar{f}(t;h_{2,n})) + o(h_{2,n}^4)$$
$$+\frac{1}{n^2h_{2,n}^2}\mathbb{E}\left[\sum_{i=1}^{n} L_g(X_i)\alpha(g(X_i))\delta_r(X_i)Y_i\sum_{i=1}^{n} K_f(X_i)\alpha(f(X_i))\right] \tag{3.2.38}$$
$$+\frac{1}{n^2h_{2,n}^2}\mathbb{E}\left[\sum_{i=1}^{n} K_g(X_i)\alpha(g(X_i))Y_i\sum_{i=1}^{n} L_f(X_i)\alpha(f(X_i))\delta(X_i)\right]. \tag{3.2.39}$$

To show that (3.2.38) is of order $o(h_{2,n}^4)$, we further decompose the term $\frac{1}{nh_{2,n}}\sum_{i=1}^{n} L_g(X_i)\alpha(g(X_i))\delta_r(X_i)Y_i$ in (3.2.38) using first the decomposition (3.1.2) of δ_r, and then the decomposition of $\hat{g}-g$ into D_2 and b_2. Then, by (3.1.5):

$$(3.2.38) = \frac{1}{n^2h_{2,n}^2}\mathbb{E}\left[\sum_{i=1}^{n} L_g(X_i)\alpha'(g(X_i))(D_2(X_i,h_{1,n})+b_2(X_i,h_{1,n}))\right.$$
$$\left. Y_i\sum_{i=1}^{n} K_f(X_i)\alpha(f(X_i))\right] + o(h_{2,n}^4)$$
$$= \mathbb{E}\left[\frac{1}{nh_{2,n}}\sum_{i=1}^{n} L_g(X_i)\alpha'(g(X_i))D_2(X_i,h_{1,n})Y_i\bar{f}(t;h_{2,n})\right] \tag{3.2.40}$$
$$+\mathbb{E}\left[\frac{1}{nh_{2,n}}\sum_{i=1}^{n} L_g(X_i)\alpha'(g(X_i))b_2(X_i,h_{1,n})Y_i\bar{f}(t;h_{2,n})\right] \tag{3.2.41}$$
$$+\,o(h_{2,n}^4).$$

We have that $\bar{f}(t;h_{2,n}) \xrightarrow{a.s.} f(t)$ and by the analysis of the term (3.22) for $d=1$ in [10] : $\frac{1}{nh_{2,n}}\sum_{i=1}^{n} L_g(X_i)\alpha'(g(X_i))b_2(X_i,h_{1,n})Y_i = o_{a.s.}(h_{2,n}^4)$. Thus, we have that,

$$(3.2.41) = o(h_{2,n}^4).$$

Estimation of (3.2.40) requires U-processes and second order Hoeffding process as in Proposition 3.2.6, specifically as in the estimation of (3.2.19). First note that the decompositions (3.2.28) – (3.2.29) of $\frac{1}{nh_{2,n}}\sum_{i=1}^{n} L_g(X_i)\alpha'(g(X_i))D_2(X_i,h_{1,n})Y_i$ in (3.2.40) is of order $o_{a.s.}(h_{2,n}^4)$ by similar analysis of the term (3.21) for one dimension as in [10], specifically (3.31) and (3.33). Next we will show that the remaining term $\mathbb{E}\left[\left(\frac{1}{nh_{1,n}h_{2,n}}\sum_{i=1}^{n}\left(\mathbb{E}_{X,Y}H_t^g\big((X,Y),(X_i,Y_i)\big)-\mathbb{E}H_t^g\right)\right)\bar{f}(t;h_{2,n})\right]$ in (3.2.40) is of order $O((nh_{2,n})^{-1})$. So,

$$\begin{aligned}
&\mathbb{E}\left[\left(\frac{1}{nh_{1,n}h_{2,n}}\sum_{i=1}^{n}\left(\mathbb{E}_{X,Y}H_t^g\big((X,Y),(X_i,Y_i)\big)\right)\bar{f}(t;h_{2,n})\right]\right.\\
&=\frac{1}{nh_{1,n}h_{2,n}^2}\mathbb{E}\left[\left(\int L_g(x)\alpha'(g(x))g(x)K\left(\frac{x-X_1}{h_{1,n}}\right)Y_1dx\right)K_f(X_1)\alpha(f(X_1))\right]\\
&+\frac{n(n-1)}{n^2h_{1,n}h_{2,n}^2}\mathbb{E}\left[\left(\int L_g(x)\alpha'(g(x))g(x)K\left(\frac{x-X_1}{h_{1,n}}\right)Y_1dx\right)\mathbb{E}K_f(X_2)\alpha(f(X_2))\right].
\end{aligned}$$

Now with the change of variables $x = uh_{1,n}+x_1$ and $x_1 = t - vh_{2,n}$ and notation (3.2.26) and similar calculations as that of (2.4.5), we have that

$$\begin{aligned}
&\mathbb{E}\left[\left(\frac{1}{nh_{1,n}h_{2,n}}\sum_{i=1}^{n}\left(\mathbb{E}_{X,Y}H_t^g\big((X,Y),(X_i,Y_i)\big)\right)\bar{f}(t;h_{2,n})\right]\right.\\
&=\left[\left(\frac{1}{h_{1,n}h_{2,n}}\mathbb{E}H_t^g\right)\mathbb{E}\bar{f}(t;h_{2,n})\right]+O((nh_{2,n})^{-1}).
\end{aligned}$$

Therefore,

$$\mathbb{E}\left[\left(\frac{1}{nh_{1,n}h_{2,n}}\sum_{i=1}^{n}\Big(\mathbb{E}_{X,Y}H_t^g\big((X,Y),(X_i,Y_i)\big)-\mathbb{E}H_t^g\Big)\right)\bar{f}(t;h_{2,n})\right]=O((nh_{2,n})^{-1}). \tag{3.2.42}$$

Recall that we have $\bar{f}(t;h_{2,n})\xrightarrow{a.s.}f(t)$ and $(3.2.28),(3.2.29)=o_{a.s.}(h_{2,n}^4)$. Then, by (3.2.42), we have that

$$(3.2.40)=o(h_{2,n}^4). \tag{3.2.43}$$

Thus, $(3.2.38)=o(h_{2,n}^4)$.

We now further decompose the term $\frac{1}{nh_{2,n}}\sum_{i=1}^{n}L_f(X_i)\alpha(f(X_i))\delta(X_i)$ in (3.2.39) using first the decomposition (2.1.7) of δ, and then the decomposition of $\hat{f}-f$ into D_1 and b_1. Then, by (2.1.6):

$$\begin{aligned}(3.2.39)=&\ \mathbb{E}\left[\frac{1}{nh_{2,n}}\sum_{i=1}^{n}L_f(X_i)\alpha'(f(X_i))D_1(X_i,h_{1,n})\bar{g}(t;h_{2,n})\right] &(3.2.44)\\ &+\mathbb{E}\left[\frac{1}{nh_{2,n}}\sum_{i=1}^{n}L_f(X_i)\alpha'(f(X_i))b_1(X_i,h_{1,n})\bar{g}(t;h_{2,n})\right] &(3.2.45)\\ &+o(h_{2,n}^4).\end{aligned}$$

Similarly, arguments from analysis of (3.2.41) and by (2.2.7), we have that $(3.2.45)=o(h_{2,n}^4)$. Also, by equivalent decomposition and work as that of (3.2.40), we have that $(3.2.44)=o(h_{2,n}^4)$.

Thus, $(3.2.40),(3.2.41)=o(h_{2,n}^4)$. Hence, $(3.2.39)=o(h_{2,n}^4)$.

Since $(3.2.38),(3.2.39)=o(h_{2,n}^4)$,

$$\mathbb{E}(\hat{g}(t;h_{1,n},h_{2,n})\hat{f}(t;h_{1,n},h_{2,n}))=\mathbb{E}(\bar{g}(t;h_{2,n})\bar{f}(t;h_{2,n}))+o(h_{2,n}^4). \tag{3.2.46}$$

Now, we work with $\mathbb{E}(\hat{g}(t;h_{1,n},h_{2,n}))\mathbb{E}(\hat{f}(t;h_{1,n},h_{2,n}))$. Using Proposition 2.2.2 from Section 2.2 for one dimension and Proposition 3.2.6, we have that

$$\begin{aligned}\mathbb{E}(\hat{f}(t;h_{1,n},h_{2,n})) &= \mathbb{E}(\bar{f}(t;h_{2,n})) + o(h_{2,n}^4),\\ \mathbb{E}(\hat{g}(t;h_{1,n},h_{2,n})) &= \mathbb{E}(\bar{g}(t;h_{2,n})) + o(h_{2,n}^4).\end{aligned}$$

Hence,

$$\begin{aligned}&\mathbb{E}(\hat{g}(t;h_{1,n},h_{2,n}))\mathbb{E}(\hat{f}(t;h_{1,n},h_{2,n}))\\ &= \mathbb{E}(\bar{g}(t,h_{2,n}))\mathbb{E}(\bar{f}(t,h_{2,n})) + o(h_{2,n}^4).\end{aligned} \tag{3.2.47}$$

Thus, by estimations (3.2.10), (3.2.46) and (3.2.47), we get that

$$\begin{aligned}&Cov\Big(\hat{g}(t;h_{1,n},h_{2,n}),\hat{f}(t;h_{1,n},h_{2,n})\Big)\\ &= \mathbb{E}(\hat{g}(t;h_{1,n},h_{2,n})\hat{f}(t;h_{1,n},h_{2,n})) - \mathbb{E}(\hat{g}(t;h_{1,n},h_{2,n}))\mathbb{E}(\hat{f}(t;h_{1,n},h_{2,n}))\\ &= Cov\Big(\bar{g}(t,h_{2,n}),\bar{f}(t,h_{2,n})\Big) + o(h_{2,n}^4) = o(h_{2,n}^4).\end{aligned}$$

Therefore, (3.2.35) $= o(h_{2,n}^4)$.

Since Y_1 is bounded, by similar argument as of $\hat{f}(t;h_{1,n},h_{2,n})$ in [10], we have that $\hat{g}(t;h_{1,n},h_{2,n}) \xrightarrow{a.s.} g(t)$. Moreover by Theorem 1 of [10] and Proposition 2.2.2, for $t \in D_{rf}$ with the condition $U(h_{1,n}) = o(h_{2,n}^2)$ and $\frac{\log n}{nh_{2,n}} = o(h_{2,n}^4)$,

$$\begin{aligned}&|\hat{f}(t;h_{1,n},h_{2,n})) - \mathbb{E}(\hat{f}(t;h_{1,n},h_{2,n}))| = |\hat{f}(t;h_{1,n},h_{2,n})) - f(t) + f(t) - \mathbb{E}(\hat{f}(t;h_{1,n},h_{2,n}))|\\ &\le |\hat{f}(t;h_{1,n},h_{2,n})) - f(t)| + |f(t) - \mathbb{E}(\hat{f}(t;h_{1,n},h_{2,n}))| = o_{\text{a.s.}}(h_{2,n}^2).\end{aligned}$$

Thus, (3.2.36) $= o(h_{2,n}^4)$.

Furthermore, for $t \in D_{rf}$, by Theorem 1 of [10] we have that $\frac{1}{\hat{f}(t;h_{1,n},h_{2,n})} \xrightarrow{a.s.} \frac{1}{f(t)}$ and $\hat{g}(t;h_{1,n},h_{2,n}) \xrightarrow{a.s.} g(t)$. Thus, we have $\hat{r}(t;h_{1,n},h_{2,n}) \xrightarrow{a.s.} r(t)$. Now by Proposition 2.2.2, we obtain

$$(3.2.37) = o(h_{2,n}^4).$$

Since $(3.2.35), (3.2.36), (3.2.37) = o(h_{2,n}^4)$, we have that

$$\mathbb{E}(\hat{r}(t; h_{1,n}, h_{2,n})) = \mathbb{E}\left(\frac{\hat{g}(t; h_{1,n}, h_{2,n})}{\hat{f}(t; h_{1,n}, h_{2,n})}\right) = \frac{\mathbb{E}(\hat{g}(t; h_{1,n}, h_{2,n}))}{\mathbb{E}(\hat{f}(t; h_{1,n}, h_{2,n}))} + o(h_{2,n}^4).$$

This proves Lemma 3.2.7. □

Theorem 3.2.8. *Define $\hat{r}(t; h_{1,n}, h_{2,n})$ by (1.4.5). Assume that $U(h_{1,n}) = o(h_{2,n}^2)$ and $\frac{\log n}{nh_{2,n}} = o(h_{2,n}^4)$. Then, under the hypotheses in Assumptions 3 with $h_n = h_{2,n}$ and Assumptions 2, for $t \in \mathcal{D}_{rf}$,*

$$\mathbb{E}(\hat{r}(t; h_{1,n}, h_{2,n})) = r(t) + \frac{\rho_4 h_{2,n}^4}{24}\left[\frac{1}{f(t)} D_4\left(\frac{1}{g(t)}\right) - \frac{g(t)}{f^2(t)} D_4\left(\frac{1}{f(t)}\right)\right] + o(h_{2,n}^4) \quad (3.2.48)$$

as $n \to \infty$.

Proof. Now, from Lemma 3.2.7 we have that

$$\mathbb{E}(\hat{r}(t; h_{1,n}, h_{2,n})) = \frac{\mathbb{E}(\hat{g}(t; h_{1,n}, h_{2,n}))}{\mathbb{E}(\hat{f}(t; h_{1,n}, h_{2,n}))} + o(h_{2,n}^4)$$

Now by Proposition 3.2.6, and Proposition 2.2.2 for the one dimension case

$$\begin{aligned}\mathbb{E}(\hat{r}(t; h_{1,n}, h_{2,n})) &= \frac{g(t) + \frac{\rho_4 h_{2,n}^4}{24}\left(D_4\left(\frac{1}{f(t)r(t)}\right)\right) + o(h_{2,n}^4)}{f(t) + \frac{\rho_4 h_{2,n}^4}{24}\left(D_4\left(\frac{1}{f(t)}\right)\right) + o(h_{2,n}^4)} \\ &= r(t) + \frac{\rho_4 h_{2,n}^4}{24}\left[\frac{1}{f(t)} D_4\left(\frac{1}{g(t)}\right) - \frac{g(t)}{f^2(t)} D_4\left(\frac{1}{f(t)}\right)\right] + o(h_{2,n}^4).\end{aligned}$$

The last equality uses the first order Taylor expansion at 0 of $\frac{1}{1+\frac{\rho_4}{24 f(t)} D_4\left(\frac{1}{f(t)}\right) h_{2,n}^4 + o(h_{2,n}^4)}$.

□

3.3 CENTRAL LIMIT THEOREM FOR VARIABLE BANDWIDTH KERNEL REGRESSION ESTIMATOR

This section presents central limit theorems for the variable bandwidth kernel regression estimators.

3.3 Ideal Estimator

Recall the notations: $K_f(X_i) := K\left(\frac{t-X_i}{h_n}\alpha(f(X_i))\right)$ and $K_g(X_i) := K\left(\frac{t-X_i}{h_n}\alpha(g(X_i))\right)$.

Theorem 3.3.1. *Let (X_i, Y_i) be i.i.d. random vectors with $1 \le i \le n$ and Y_1 be bounded. Define $\bar{r}(t; h_n)$ by equation (3.2.1). Then, under Assumptions 3 for $t \in D_{rf}$,*

$$\sqrt{nh_n}[\bar{r}(t; h_n) - r(t)] \xrightarrow{D} N\left(\frac{\lambda q(t)}{f^2(t)}, \sigma_r^2\right), \tag{3.3.1}$$

where $h_n^4\sqrt{nh_n} \to \lambda$ with $0 \le \lambda < \infty$ as $n \to \infty$.
Here $\sigma_r^2 := \frac{\sigma_u^2}{f^4(t)}$,

where $\sigma_u^2 = f^3(t)m(t)\alpha(g(t))\mu_0 + g^2(t)f(t)\alpha(f(t))\mu_0 - 2f(t)g(t)\beta(t)$,
$\beta(t) = g(t)\alpha(f(t))\alpha(g(t))\int K(u\alpha(g(t)))\, K(u\alpha(f(t)))\, du$, $m(x) = \int y^2 f(y|x)dy$, $q(t) := \frac{\rho_4}{24}\left[f(t)D_4\left(\frac{1}{g(t)}\right) - g(t)D_4\left(\frac{1}{f(t)}\right)\right]$, $\mu_0 = \int K^2(w)dw$ and $\rho_k = \int w^k K(w)dw$.

The following lemma is necessary in the proof of Theorem 3.3.1.

Lemma 3.3.2. *Let (X_i, Y_i) be i.i.d. random vectors with $1 \le i \le n$ and Y_1 be bounded. Define $\bar{g}(t; h_n)$ by equation (3.2.1). Then, under Assumptions 3, as $n \to \infty$,*

$$Var(\bar{g}(t; h_n)) = \frac{m(t)f(t)\alpha(g(t))\mu_0}{nh_n}(1 + o(1)). \tag{3.3.2}$$

Proof. Let $h_n = h$. Consider the second moment of $\bar{g}(t; h)$.

Denote $T(X_i) = \alpha(g(X_i))K(h^{-1}\alpha(g(X_i))(t - X_i))Y_i$. Then,

$$\begin{aligned}
\mathbb{E}(\bar{g}^2(t; h)) &= \frac{1}{n^2h^2}\mathbb{E}\left(\sum_{i=1}^{n} T(X_i)\right)^2 \\
&= \frac{1}{n^2h^2}\sum_{i=1}^{n}\mathbb{E}T^2(X_i) + \frac{1}{n^2h^2}\sum_{i\neq j}\mathbb{E}T(X_i)\mathbb{E}T(X_j) \\
&= \frac{1}{nh^2}\mathbb{E}T^2(X_1) + \frac{n(n-1)}{n^2h^2}(\mathbb{E}T(X_1))^2.
\end{aligned}$$

Recall that $g(t) = f(t)r(t)$. Thus, by (3.2.4), $\mathbb{E}(\bar{g}(t; h)) = g(t) + O(h^2)$. Taking into account that $\mathbb{E}T(X_1) = h\mathbb{E}(\bar{g}(t; h))$, we obtain

$$\begin{aligned}
Var(\bar{g}(t; h)) &= \mathbb{E}(\bar{g}^2(t; h)) - [\mathbb{E}(\bar{g}(t; h))]^2 \\
&= \frac{1}{nh^2}\mathbb{E}T^2(X_1) + \frac{(n-1)}{nh^2}(\mathbb{E}T(X_1))^2 - \frac{1}{h^2}(\mathbb{E}T(X_1))^2 \\
&= \frac{1}{nh^2}\mathbb{E}T^2(X_1) - \frac{1}{n}\left[g(t) + O(h^2)\right]^2 \\
&= \frac{1}{nh^2}\mathbb{E}T^2(X_1) + O(n^{-1}). \qquad (3.3.3)
\end{aligned}$$

By Proposition 2.3.2 with $\zeta(s) = f(s)m(s)$ and $d = 1$, we have

$$\mathbb{E}T^2(X_1) = h\alpha(g(t))m(t)f(t)\mu_0 + o(h). \qquad (3.3.4)$$

Thus, the variance of the term $\bar{g}(t; h)$ is $\frac{m(t)f(t)\alpha(g(t))\mu_0}{nh}(1 + o(1))$. □

Proof of Theorem 3.3.1. Now, consider the following decomposition

$$\begin{aligned}
[\bar{r}(t; h_n) - r(t)] &= \frac{\bar{g}(t; h_n)}{\bar{f}(t; h_n)} - \frac{g(t)}{f(t)} = \frac{\bar{g}(t; h_n)f(t) - g(t)\bar{f}(t; h_n)}{f(t)\bar{f}(t; h_n)} \\
&= \frac{\frac{1}{n}\sum_{i=1}^{n} W_{n,i}}{f(t)\bar{f}(t; h_n)} = \frac{\bar{W}_n}{f(t)\bar{f}(t; h_n)}.
\end{aligned}$$

Define:

$$W_{n,i} = \frac{1}{h_n}\Big(f(t)Y_iK_g(X_i)\alpha(g(X_i)) - g(t)K_f(X_i)\alpha(f(X_i))\Big). \tag{3.3.5}$$

Then, by Corollary 2.2.1 and (3.2.4), for $t \in D_{rf}$,

$$\begin{aligned}
\mathbb{E}(W_{n,1}) &= f(t)\left[g(t) + \frac{\rho_4 h_n^4}{24}D_4\left(\frac{g(t)}{\alpha^4(g(t))}\right) + o(h_n^4)\right] \\
&\quad - g(t)\left[f(t) + \frac{\rho_4 h_n^4}{24}D_4\left(\frac{f(t)}{\alpha^4(f(t))}\right) + o(h_n^4)\right] \\
&= \frac{h_n^4\rho_4}{24}\Big[f(t)D_4\left(\frac{1}{g(t)}\right) - g(t)D_4\left(\frac{1}{f(t)}\right)\Big] + o(h_n^4) \\
&= h_n^4 q(t) + o(h_n^4). \qquad (3.3.6)
\end{aligned}$$

Now, we shall work on the second moment of $W_{n,1}$ to find the variance of $W_{n,1}$.

$$\begin{aligned}
\mathbb{E}(W_{n,1}^2) &= \frac{1}{h_n^2}\mathbb{E}\Big(f(t)Y_1K_g(X_1)\alpha(g(X_1)) - g(t)K_f(X_1)\alpha(f(X_1))\Big)^2 \\
&= \frac{f^2(t)}{h_n^2}\mathbb{E}\Big(Y_1^2K_g^2(X_1)\alpha^2(g(X_1))\Big) + \frac{g^2(t)}{h_n^2}\mathbb{E}\left(K_f^2(X_1)\alpha^2(f(X_1))\right) \\
&\quad - \frac{2f(t)g(t)}{h_n^2}\mathbb{E}\left(Y_1K_g(X_1)\alpha(g(X_1))K_f(X_1)\alpha(f(X_1))\right) \\
&= \frac{f^2(t)}{h_n^2}\mathbb{E}\Big(m(X_1)K_g^2(X_1)\alpha^2(g(X_1))\Big) + \frac{g^2(t)}{h_n^2}\mathbb{E}\left(K_f^2(X_1)\alpha^2(f(X_1))\right) \\
&\quad - \frac{2f(t)g(t)}{h_n^2}\mathbb{E}\left(r(X_1)K_g(X_1)\alpha(g(X_1))K_f(X_1)\alpha(f(X_1))\right). \qquad (3.3.7)
\end{aligned}$$

Then, by Proposition 2.3.2 for $d = 1$, particularly from (2.3.8), we have

$$\mathbb{E}\Big(m(X_1)K_g^2(X_1)\alpha^2(g(X_1))\Big) = (1 + o(1))(h_n m(t)f(t)\alpha(g(t))\mu_0),$$

by Lemma 3.3.2, we have

$$\mathbb{E}\Bigg(K_f^2(X_1)\alpha^2(f(X_1))\Bigg) = (1+o(1))(h_n\alpha(f(t))f(t)\mu_0),$$

and by change of variable: $X_1 = t - h_n u$, particularly (3.2.12), we have

$$(3.3.7) = \frac{2f(t)g(t)}{h_n}\beta_n(t).$$

Thus,

$$\begin{aligned}\mathbb{E}(W_{n,1}^2) = &\left(\frac{f^2(t)}{h_n}\Big(m(t)f(t)\alpha(g(t))\mu_0\Big) + \frac{g^2(t)}{h_n}\Big(\alpha(f(t))f(t)\mu_0\Big)\right)(1+o(1)) \\ &- \frac{2f(t)g(t)}{h_n}\beta_n(t),\end{aligned}$$

$$\begin{aligned}\text{where } \beta_n(t) := \int &K\left(u\alpha(g(t-h_nu))\right)\alpha(g(t-h_nu)) \\ &K\left(u\alpha(f(t-h_nu))\right)\alpha(f(t-h_nu))g(t-h_nu)du.\end{aligned}$$

Hence,

$$\begin{aligned}Var(W_{n,1}) &= \mathbb{E}(W_{n,1}^2) - (\mathbb{E}(W_{n,1}))^2 \\ &= \left(\frac{f^2(t)}{h_n}m(t)f(t)\alpha(g(t))\mu_0 + \frac{g^2(t)}{h_n}\alpha(f(t))f(t)\mu_0\right)(1+o(1)) \\ &\quad - \frac{2f(t)g(t)}{h_n}\beta_n(t) - h_n^8 q^2(t) + o(h_n^8).\end{aligned}$$

Thus, by Theorem 2.4.2,

$$\sqrt{nh_n}[\bar{W}_n - \mathbb{E}(W_{n,1})] \xrightarrow{D} N\left(0, \sigma_u^2\right), \tag{3.3.8}$$

where $\sigma_u^2 = f^3(t)m(t)\alpha(g(t))\mu_0 + g^2(t)f(t)\alpha(f(t))\mu_0 - 2f(t)g(t)\beta(t)$.

Note that when K, α, g, and f are bounded and K has compact support,

$$\begin{aligned}\beta_n(t) = \int & K\left(u\alpha(g(t-h_nu))\right)\alpha(g(t-h_nu)) \\ & K\left(u\alpha(f(t-h_nu))\right)\alpha(f(t-h_nu))g(t-h_nu)du \\ & \xrightarrow{n\to\infty} g(t)\alpha(f(t))\alpha(g(t))\int K\left(u\alpha(g(t))\right)K\left(u\alpha(f(t))\right)du = \beta(t)\end{aligned}$$

For $t \in D_{rf}$, $f(t) > 0$ and recall that by Corollary 2.2.1 and Lemma 3.2.2 we have that $\frac{1}{\bar{f}(t;h_n)} \xrightarrow{a.s} \frac{1}{f(t)}$. Assuming $h_n^4\sqrt{nh_n} \to \lambda$, using Theorem 2.4.1, (3.3.8) and convergence of $1/\bar{f}$ a.s., we have

$$\sqrt{nh_n}[\bar{r}(t;h_n) - r(t)] \xrightarrow{D} N\left(\frac{\lambda q(t)}{f^2(t)}, \sigma_r^2\right).$$

$\square$

3.3 True Estimator

Theorem 3.3.3. *Let (X_i, Y_i) be i.i.d. random vectors with $1 \le i \le n$ and Y_1 be bounded. Define $\hat{r}(t;h_{1,n},h_{2,n})$ by equation (1.4.5). Assume $U(h_{1,n}) = o(h_{2,n}^2)$ and $\frac{\log n}{nh_{2,n}} = o(h_{2,n}^4)$. Then, under Assumptions 2 and Assumptions 3 with $h_n = h_{2,n}$, for $t \in D_{rf}$, as $n \to \infty$,*

$$\sqrt{nh_{2,n}}[\hat{r}(t;h_{1,n},h_{2,n}) - r(t)] \xrightarrow{D} N\left(\frac{\lambda q(t)}{f^2(t)}, \sigma_{r1}^2\right), \tag{3.3.9}$$

where $h_{2,n}^4\sqrt{nh_{2,n}} \to \lambda$ with $0 \le \lambda < \infty$ as $n \to \infty$, and

$$
\begin{aligned}
\sigma_{r1}^2 = \sigma_r^2 + r^2(t)\Bigg[& f(t)\frac{[\alpha'(f(t))]^2}{\alpha(f(t))}\int L^2(z)dz + f(t)m(t)\frac{[\alpha'(g(t))]^2}{\alpha(g(t))}\int L^2(z)dz \\
& - 2g(t)\alpha'(g(t))\alpha'(f(t))\int L(v\alpha(f(t)))L(v\alpha(g(t)))dv\Bigg] \\
& + r^2(t)\Bigg[\alpha'(f(t))\mu_0 + \alpha'(g(t))\mu_0 \\
& - 2r(t)\alpha'(g(t))\alpha(f(t))\int L(v\alpha(g(t)))K(v\alpha(f(t)))dv \\
& - 2\alpha'(f(t))\alpha(g(t))\int L(v\alpha(f(t)))K(v\alpha(g(t)))dv\Bigg].
\end{aligned}
$$

$q(t) := \frac{\rho_4}{24}\left[f(t)D_4\left(\frac{1}{g(t)}\right) - g(t)D_4\left(\frac{1}{f(t)}\right)\right]$, *$\rho_k = \int w^k K(w)dw$, $m(x) = \int y^2 f(y|x)dy$, $\mu_0 = \int K^2(w)dw$, and $L(t) = tK'(t) + K(t)$.*

Under the notations in Theorem 3.3.3, let

$$
\begin{aligned}
\sigma_{u1}^2 = \sigma_u^2 + g^2(t)f^2(t)\Bigg[& f(t)\frac{[\alpha'(f(t))]^2}{\alpha(f(t))}\int L^2(z)dz + f(t)m(t)\frac{[\alpha'(g(t))]^2}{\alpha(g(t))}\int L^2(z)dz \\
& - 2g(t)\alpha'(g(t))\alpha'(f(t))\int L(v\alpha(f(t)))L(v\alpha(g(t)))dv\Bigg] \\
& + f(t)g^2(t)\Bigg[f(t)\alpha'(f(t))\mu_0 + f(t)\alpha'(g(t))\mu_0 \\
& - 2g(t)\alpha'(g(t))\alpha(f(t))\int L(v\alpha(g(t)))K(v\alpha(f(t)))dv \\
& - 2f(t)\alpha'(f(t))\alpha(g(t))\int L(v\alpha(f(t)))K(v\alpha(g(t)))dv\Bigg]. \qquad (3.3.10)
\end{aligned}
$$

The results from the following lemma is necessary for the proof of Theorem 3.3.3.

Lemma 3.3.4. *Under the conditions and notations in Theorem 3.3.3, we have*

$$
\mathbb{E}(T_{n,1}) = h_{2,n}^4 q(t) + o(h_{2,n}^4),
$$

$$h_{2,n}Var(T_{n,1}) \xrightarrow{n\to\infty} \sigma_{u1}^2,$$

where σ_{u1}^2 is defined in (3.3.10), $T_{n,1} = W_{n,1} + P_{n,1}$, $P_{n,1} = \frac{f(t)}{h_{1,n}h_{2,n}}\Big(\mathbb{E}_{X,Y}H_t^g\big((X,Y),(X_1,Y_1)\big) - \mathbb{E}H_t^g\Big) - \frac{g(t)}{h_{1,n}h_{2,n}}(\mathbb{E}_X H_t(X,X_1) - \mathbb{E}H_t)$, $W_{n,1}$ is defined as (3.3.5), H_t and H_t^g is defined as (2.2.10) and (3.2.26) respectively.

Proof. First, we work on expectation of $T_{n,1}$. Notice that $\mathbb{E}(P_{n,1}) = 0$. Thus, by (3.3.6) we get

$$\mathbb{E}(T_{n,1}) = \mathbb{E}(W_{n,1}) + \mathbb{E}(P_{n,1}) = \mathbb{E}(W_{n,1}) = h_{2,n}^4 q(t) + o(h_{2,n}^4). \tag{3.3.11}$$

We will work on the variance of $T_{n,1}$ next. By the definition of variance

$$\begin{aligned} h_{2,n}Var(T_{n,1}) &= h_{2,n}[\mathbb{E}T_{n,1}^2 - (\mathbb{E}T_{n,1})^2] \\ &= h_{2,n}\mathbb{E}W_{n,1}^2 + h_{2,n}\mathbb{E}P_{n,1}^2 + 2h_{2,n}\mathbb{E}(P_{n,1}W_{n,1}) - h_{2,n}(\mathbb{E}T_{n,1})^2. \end{aligned}$$

By (3.3.7), note that the limiting term of $h_{2,n}\mathbb{E}W_{n,1}^2$ is σ_u^2. We only need to calculate the limit of terms $h_{2,n}\mathbb{E}(P_{n,1}W_{n,1})$ and $h_{2,n}\mathbb{E}P_{n,1}^2$. First we will find the limit of the term $h_{2,n}\mathbb{E}(P_{n,1}W_{n,1})$.

$$\begin{aligned} P_{n,1}W_{n,1} &= \left[\frac{f(t)}{h_{1,n}h_{2,n}}\Big(\mathbb{E}_{X,Y}H_t^g\big((X,Y),(X_1,Y_1)\big) - \mathbb{E}H_t^g\Big) - \frac{g(t)}{h_{1,n}h_{2,n}}(\mathbb{E}_X H_t(X,X_1) - \mathbb{E}H_t)\right] \\ &\times \left[\frac{f(t)}{h_{2,n}}Y_1 K_g(X_1)\alpha(g(X_1)) - \frac{g(t)}{h_{2,n}}K_f(X_1)\alpha(f(X_1))\right] \\ &= \frac{f^2(t)}{h_{1,n}h_{2,n}^2}\Big(\mathbb{E}_{X,Y}H_t^g\big((X,Y),(X_1,Y_1)\big) - \mathbb{E}H_t^g\Big)K_g(X_1)\alpha(g(X_1))Y_1 \qquad (3.3.12) \\ &+ \frac{g^2(t)}{h_{1,n}h_{2,n}^2}(\mathbb{E}_X H_t(X,X_1) - \mathbb{E}H_t)K_f(X_1)\alpha(f(X_1)) \qquad (3.3.13) \\ &- \frac{f(t)g(t)}{h_{1,n}h_{2,n}^2}(\mathbb{E}_{X,Y}H_t^g\big((X,Y),(X_1,Y_1) - \mathbb{E}H_t^g\big)K_f(X_1)\alpha(f(X_1)) \qquad (3.3.14) \end{aligned}$$

$$-\frac{f(t)g(t)}{h_{1,n}h_{2,n}^2}(\mathbb{E}_X H_t(X,X_1)-\mathbb{E}H_t)K_g(X_1)\alpha(g(X_1))Y_1. \tag{3.3.15}$$

From Section 2.4, namely (2.4.7) with $d=1$, we have that

$$2h_{2,n}\mathbb{E}(3.3.13) \xrightarrow{n\to\infty} g^2(t)f^2(t)\alpha'(f(t))\mu_0. \tag{3.3.16}$$

Also note that by the change of variables $x_1 = x - uh_{1,n}$ and $x_1 = t - vh_{2,n}$, $\mathbb{E}(H_t)$ and $\mathbb{E}(H_t^g)$ are bounded by $Ch_{1,n}h_{2,n}$ for some $C>0$ and then

$$\frac{1}{h_{1,n}}\mathbb{E}(H_t) \to 0 \text{ and } \frac{1}{h_{1,n}}\mathbb{E}(H_t^g) \to 0 \text{ as } n\to\infty. \tag{3.3.17}$$

For the remaining terms in $h_{2,n}\mathbb{E}(P_{n,1}W_{n,1})$, with the same change of variables $x_1 = x - uh_{1,n}$ and $x_1 = t - vh_{2,n}$, we have

$$\begin{aligned}
2h_{2,n}\mathbb{E}(3.3.12) = {} & \frac{2f^2(t)}{h_{1,n}h_{2,n}}\int\int L_g(x)\alpha'(g(x))K\left(\frac{x-x_1}{h_{1,n}}\right)K_g(x_1) \\
& \times \alpha(g(x_1))g(x)g(x_1)dxdx_1 \\
& - \frac{2f^2(t)}{h_{1,n}h_{2,n}}\mathbb{E}H_t^g\mathbb{E}\Big(K_g(X_1)\alpha(g(X_1))Y_1\Big) \\
= 2f^2(t)\int\int L & \left(\left(v-\frac{uh_{1,n}}{h_{2,n}}\right)\alpha(g(uh_{1,n}+t-vh_{2,n}))\right) \\
& \times \alpha'(g(uh_{1,n}+t-vh_{2,n}))K(u)K(v\alpha(g(t-h_{2,n}v))) \\
& \times \alpha(g(t-h_{2,n}v))g(t-h_{2,n}v)g(uh_{1,n}+t-vh_{2,n})dudv \\
& - \frac{2f^2(t)}{h_{1,n}}\mathbb{E}H_t^g\int K(v\alpha(g(t-h_{2,n}v)))\alpha(g(t-h_{2,n}v))g(t-h_{2,n}v)dv.
\end{aligned}$$

Since K, α, g , L are bounded and K has compact support, by (3.3.17),

$$2h_{2,n}\mathbb{E}(3.3.12) \xrightarrow{n\to\infty} f^2(t)g^2(t)\alpha'(g(t))\mu_0. \tag{3.3.18}$$

For the following remaining terms in $h_{2,n}\mathbb{E}(P_{n,1}W_{n,1})$, we use (3.3.17) and the fact that K, α, g , L are bounded and that K has compact support.

$$
\begin{aligned}
2h_{2,n}\mathbb{E}(3.3.14) = &\frac{2f(t)g(t)}{h_{1,n}h_{2,n}} \int\int L_g(x)\alpha'(g(x))K\left(\frac{x-x_1}{h_{1,n}}\right)K_f(x_1) \\
&\times \alpha(f(x_1))g(x)g(x_1)dxdx_1 \\
&- \frac{2f(t)g(t)}{h_{1,n}h_{2,n}}\mathbb{E}H_t^g\mathbb{E}\big(K_f(X_1)\alpha(f(X_1))\big) \\
= 2f(t)g(t)\int\int L&\left(\left(v-\frac{uh_{1,n}}{h_{2,n}}\right)\alpha(g(uh_{1,n}+t-vh_{2,n}))\right) \\
&\times \alpha'(g(uh_{1,n}+t-vh_{2,n}))K(u)K(v\alpha(f(t-h_{2,n}v))) \\
&\times \alpha(f(t-h_{2,n}v))g(t-h_{2,n}v)g(uh_{1,n}+t-vh_{2,n})dudv \\
- \frac{2f(t)g(t)}{h_{1,n}}\mathbb{E}H_t^g &\int K(v\alpha(f(t-h_{2,n}v)))\alpha(f(t-h_{2,n}v))f(t-h_{2,n}v)dv \\
\xrightarrow{n\to\infty} 2f(t)g^3(t)&\alpha'(g(t))\alpha(f(t))\int L(v\alpha(g(t)))K(v\alpha(f(t)))dv. \qquad (3.3.19)
\end{aligned}
$$

$$
\begin{aligned}
2h_{2,n}\mathbb{E}(3.3.15) = &\frac{2f(t)g(t)}{h_{1,n}h_{2,n}} \int\int L_f(x)\alpha'(f(x))K\left(\frac{x-x_1}{h_{1,n}}\right)K_g(x_1) \\
&\times \alpha(g(x_1))f(x)g(x_1)dxdx_1 \\
&- \frac{2f(t)g(t)}{h_{1,n}h_{2,n}}\mathbb{E}H_t\mathbb{E}\Big(K_g(X_1)\alpha(g(X_1))Y_1\Big) \\
= 2f(t)g(t)\int\int L&\left(\left(v-\frac{uh_{1,n}}{h_{2,n}}\right)\alpha(f(uh_{1,n}+t-vh_{2,n}))\right) \\
&\alpha'(f(uh_{1,n}+t-vh_{2,n}))K(u)K(v\alpha(g(t-h_{2,n}v))) \\
&\times \alpha(g(t-h_{2,n}v))g(t-h_{2,n}v)f(uh_{1,n}+t-vh_{2,n})dudv \\
&- \frac{2f(t)g(t)}{h_{1,n}}\mathbb{E}H_t\int K(v\alpha(g(t-h_{2,n}v)))\alpha(g(t-h_{2,n}v))g(t-h_{2,n}v)dv \\
\xrightarrow{n\to\infty} 2f^2(t)g^2(t)&\alpha'(f(t))\alpha(g(t))\int L(v\alpha(f(t)))K(v\alpha(g(t)))dv. \qquad (3.3.20)
\end{aligned}
$$

So, putting (3.3.16) – (3.3.20) together, we get:

$$2h_{2,n}\mathbb{E}(P_{n,1}W_{n,1}) \xrightarrow{n\to\infty} f(t)g^2(t)\Big(f(t)\alpha'(f(t))\mu_0 + f(t)\alpha'(g(t))\mu_0 - 2g(t)\alpha'(g(t))\alpha(f(t)) \int L(v\alpha(g(t)))K(v\alpha(f(t)))dv - 2f(t)\alpha'(f(t))\alpha(g(t))\int L(v\alpha(f(t)))K(v\alpha(g(t)))dv\Big). \tag{3.3.21}$$

Next we will find the limit of the term $h_{2,n}\mathbb{E}P_{n,1}^2$.

$$\begin{aligned} P_{n,1}^2 &= \left[\frac{f(t)}{h_{1,n}h_{2,n}}\Big(\mathbb{E}_{X,Y}H_t^g\big((X,Y),(X_1,Y_1)\big) - \mathbb{E}H_t^g\Big) - \frac{g(t)}{h_{1,n}h_{2,n}}(\mathbb{E}_X H_t(X,X_1) - \mathbb{E}H_t)\right]^2 \\ &= \frac{f^2(t)}{h_{1,n}^2h_{2,n}^2}\Big(\mathbb{E}_{X,Y}H_t^g\big((X,Y),(X_1,Y_1)\big) - \mathbb{E}H_t^g\Big)^2 && (3.3.22) \\ &+ \frac{g^2(t)}{h_{1,n}^2h_{2,n}^2}(\mathbb{E}_X H_t(X,X_1) - \mathbb{E}H_t)^2 && (3.3.23) \\ &- \frac{2f(t)g(t)}{h_{1,n}^2h_{2,n}^2}\Big(\mathbb{E}_{X,Y}H_t^g\big((X,Y),(X_1,Y_1)\big) - \mathbb{E}H_t^g\Big)(\mathbb{E}_X H_t(X,X_1) - \mathbb{E}H_t). && (3.3.24) \end{aligned}$$

Again, by (2.4.6) for $d = 1$ in Section 2.4

$$h_{2,n}\mathbb{E}(3.3.23) \xrightarrow{n\to\infty} f^3(t)g^2(t)\frac{[\alpha'(f(t))]^2}{\alpha(f(t))}\int L^2(z)dz. \tag{3.3.25}$$

For the remaining terms in $h_{2,n}\mathbb{E}(P_{n,1}^2)$ with the change of variables $x_1 = x - uh_{1,n}$ and $x_1 = t - vh_{2,n}$, we will calculate the limit for the following term, particularly for the term (3.3.22) :

$$\begin{aligned} &\frac{f^2(t)}{h_{1,n}^2h_{2,n}}\mathbb{E}\Big(\mathbb{E}_{X,Y}H_t^g\big((X,Y),(X_1,Y_1)\big)\Big)^2 \\ &= \frac{f^2(t)}{h_{1,n}^2h_{2,n}}\int\left[\int L_g(x)\alpha'(g(x))K\left(\frac{x-x_1}{h_{1,n}}\right)g(x)dx\right]^2 m(x_1)f(x_1)dx_1 \end{aligned}$$

$$
\begin{aligned}
&= f^2(t)\int\Big[\int L\left(\left(v-\frac{h_{1,n}u}{h_{2,n}}\right)\alpha(g(t-h_{2,n}v+h_{1,n}u))\right)\\
&\qquad\times\alpha'(g(t-h_{2,n}v+h_{1,n}u))K(u)g(t-h_{2,n}v+h_{1,n}u)du\Big]^2\\
&\qquad\times m(t-h_{2,n}v)f(t-h_{2,n}v)dv
\end{aligned}
$$

Then, by taking the limit, we get

$$
\frac{f^2(t)}{h_{1,n}^2h_{2,n}}\mathbb{E}\Big(\mathbb{E}_{X,Y}H_t^g\big((X,Y),(X_1,Y_1)\big)\Big)^2 \xrightarrow{n\to\infty} f^3(t)g^2(t)m(t)\frac{[\alpha'(g(t))]^2}{\alpha(g(t))}\int L^2(z)dz. \tag{3.3.26}
$$

Thus, by (3.3.17), (3.3.26) and boundedness of K, α, g , L with the compact support of K:

$$
h_{2,n}\mathbb{E}(3.3.22) \xrightarrow{n\to\infty} f^3(t)g^2(t)m(t)\frac{[\alpha'(g(t))]^2}{\alpha(g(t))}\int L^2(z)dz. \tag{3.3.27}
$$

To work on the expectation of the term (3.3.24), we first calculate the following term:

$$
\begin{aligned}
&\frac{2f(t)g(t)}{h_{1,n}^2h_{2,n}}\mathbb{E}\Big(\mathbb{E}_{X,Y}H_t^g\big((X,Y),(X_1,Y_1)\big)\Big)(\mathbb{E}_XH_t(X,X_1))\\
&=\frac{2f(t)g(t)}{h_{1,n}^2h_{2,n}}\int\left[\int L_g(x)\alpha'(g(x))K\left(\frac{x-x_1}{h_{1,n}}\right)g(x)dx\right]\\
&\qquad\qquad\left[\int L_f(x)\alpha'(f(x))K\left(\frac{x-x_1}{h_{1,n}}\right)f(x)dx\right]g(x_1)dx_1\\
&=2f(t)g(t)\int\Big[\int L\left(\left(v-\frac{h_{1,n}u}{h_{2,n}}\right)\alpha(g(t-h_{2,n}v+h_{1,n}u))\right)\\
&\times\alpha'(g(t-h_{2,n}v+h_{1,n}u))K(u)g(t-h_{2,n}v+h_{1,n}u)du\Big]\\
&\times\Big[\int L\left(\left(v-\frac{h_{1,n}u}{h_{2,n}}\right)\alpha(f(t-h_{2,n}v+h_{1,n}u))\right)\\
&\times\alpha'(f(t-h_{2,n}v+h_{1,n}u))K(u)f(t-h_{2,n}v+h_{1,n}u)du\Big]
\end{aligned}
$$

$$\times g(t-h_{2,n}v)dv$$
$$\xrightarrow{n\to\infty} 2g^3(t)f^2(t)\alpha'(g(t))\alpha'(f(t))\int L(v\alpha(f(t)))L(v\alpha(g(t)))dv. \tag{3.3.28}$$

By (3.3.17), (3.3.28) and boundedness of K, α, g , L with the compact support of K:

$$h_{2,n}\mathbb{E}(3.3.24) \xrightarrow{n\to\infty} 2g^3(t)f^2(t)\alpha'(g(t))\alpha'(f(t))\int L(v\alpha(f(t)))L(v\alpha(g(t)))dv. \tag{3.3.29}$$

So, as a consequence of the limit of the terms in $h_{2,n}\mathbb{E}P^2_{n,1}$, (3.3.25) $-$ (3.3.29), we get:

$$\begin{aligned}h_{2,n}\mathbb{E}P^2_{n,1} \xrightarrow{n\to\infty} f^2(t)g^2(t)\Bigg[f(t)\frac{[\alpha'(f(t))]^2}{\alpha(f(t))}\int L^2(z)dz\\ + f(t)m(t)\frac{[\alpha'(g(t))]^2}{\alpha(g(t))}\int L^2(z)dz\\ - 2g(t)\alpha'(g(t))\alpha'(f(t))\int L(v\alpha(f(t)))L(v\alpha(g(t)))dv\Bigg].\end{aligned} \tag{3.3.30}$$

Thus, by (3.3.21) and (3.3.30),

$$h_{2,n}Var(T_{n,1}) \xrightarrow{n\to\infty} \sigma^2_{u1} \tag{3.3.31}$$

$\square$

Proof of Theorem 3.3.3. Now, consider the following decomposition

$$\begin{aligned}[\hat{r}(t;h_{1,n},h_{2,n})-r(t)] &= \frac{\hat{g}(t;h_{1,n}h_{2,n})}{\hat{f}(t;h_{1,n}h_{2,n})} - \frac{g(t)}{f(t)} = \frac{\hat{g}(t;h_{1,n}h_{2,n})f(t)-g(t)\hat{f}(t;h_{1,n}h_{2,n})}{f(t)\hat{f}(t;h_{1,n}h_{2,n})}\\ &= \frac{\frac{1}{n}\sum_{i=1}^n \hat{W}_{n,i}}{f(t)\hat{f}(t;h_{1,n}h_{2,n})},\end{aligned}$$

where

$$\hat{W}_{n,i} = \frac{1}{h_{2,n}}\Big(f(t)Y_i K\left(\frac{t-X_i}{h_{2,n}}\alpha(\hat{g}(X_i;h_{1,n}))\right)\alpha(\hat{g}(X_i;h_{1,n})) - g(t)$$
$$K\left(\frac{t-X_i}{h_{2,n}}\alpha(\hat{f}(X_i;h_{1,n}))\right)\alpha(\hat{f}(X_i;h_{1,n}))\Big).$$

Then, decompositions (2.1.2), (3.1.1), and the Taylor expansion of $K\left(\frac{t-X_i}{h_{2,n}}\alpha(\hat{g}(X_i;h_{1,}))\right)$ in (3.1.7), and $K\left(\frac{t-X_i}{h_{2,n}}\alpha(\hat{f}(X_i;h_{1,n}))\right)$ in (2.1.10) give,

$$\frac{1}{n}\sum_{i=1}^{n}\hat{W}_{n,i} = \frac{1}{n}\sum_{i=1}^{n}W_{n,i}$$
$$+\frac{1}{nh_{2,n}}\sum_{i=1}^{n}\Big[f(t)L_g(X_i)\alpha(g(X_i))\delta_r(X_i)Y_i - g(t)L_f(X_i)\alpha(f(X_i))\delta(X_i)\Big] \quad (3.3.32)$$
$$+\frac{1}{nh_{2,n}}\sum_{i=1}^{n}\Big[f(t)L_1\left(\frac{t-X_i}{h_{2,n}}\alpha(g(X_i))\right)\alpha(g(X_i))\delta_r^2(X_i)Y_i$$
$$- g(t)L_1\left(\frac{t-X_i}{h_{2,n}}\alpha(f(X_i))\right)\alpha(f(X_i))\delta^2(X_i)\Big] \quad (3.3.33)$$
$$+\frac{1}{nh_{2,n}}\sum_{i=1}^{n}\Big[f(t)\alpha(g(X_i))\delta_3(t,X_i)\delta_r(X_i)Y_i - g(t)\alpha(f(X_i))\delta_2(t,X_i)\delta(X_i)\Big] \quad (3.3.34)$$
$$+\frac{1}{nh_{2,n}}\sum_{i=1}^{n}\Big[f(t)\alpha(g(X_i))\delta_3(t,X_i)Y_i - g(t)\alpha(f(X_i))\delta_2(t,X_i)\Big]. \quad (3.3.35)$$

By similar analysis for (3.20) of [10], particularly Lemma 2.2.3, (2.1.6), (2.1.12), (3.1.5), and (3.1.8) and the boundedness of α, K, Y_1 and L_1, we have

$$\sup_{t\in\mathbb{R}}\left|\Big((3.3.33)+(3.3.34)+(3.3.35)\Big)\right| = O_{a.s}(U^2(h_{1,n})) = o_{a.s}(h_{2,n}^4). \quad (3.3.36)$$

We further decompose (3.3.32) using the expansion (2.1.7) and (3.1.2) of δ and δ_r respectively, and the decomposition of $\hat{f} - f$ into D_1 and b_1 and $\hat{g} - g$ into D_2 and b_2 respectively.

$$\begin{aligned}
(3.3.32) = \frac{1}{nh_{2,n}} \sum_{i=1}^{n} \Bigg[& f(t) L_g(X_i) \alpha'(g(X_i)) D_2(X_i; h_{1,n}) \\
& - g(t) L_f(X_i) \alpha'(f(X_i)) D_1(X_i; h_{1,n}) \Bigg] \qquad (3.3.37) \\
+ \frac{1}{nh_{2,n}} \sum_{i=1}^{n} \Bigg[& f(t) L_g(X_i) \alpha'(g(X_i)) b_2(X_i; h_{1,n}) \\
& - g(t) L_f(X_i) \alpha'(f(X_i)) b_1(X_i; h_{1,n}) \Bigg] \qquad (3.3.38) \\
+ \Bigg[\frac{f(t)}{2nh_{2,n}} \sum_{i=1}^{n} & \left(L_g(X_i) (\alpha)''(\gamma(X_i)) [\hat{g}(X_i; h_{1,n}) - g(X_i)]^2 \right) \\
- \frac{g(t)}{2nh_{2,n}} \sum_{i=1}^{n} & \left(L_f(X_i) (\alpha)''(\eta(X_i)) [\hat{f}(X_i; h_{1,n}) - f(X_i)]^2 \right) \Bigg]. \qquad (3.3.39)
\end{aligned}$$

Then, by (2.1.6) and (3.1.5)

$$\sup_{t \in \mathbb{R}} |(3.3.39)| = O_{a.s} \left(\frac{U^2(h_{1,n})}{h_{2,n}} \right) = o_{a.s}(h_{2,n}^4). \qquad (3.3.40)$$

Also, since Y_1 is bounded, by similar argument as (3.27) in [10],

$$\sup_{t \in \mathbb{R}} |(3.3.38)| = o_{a.s}(h_{2,n}^4). \qquad (3.3.41)$$

We decompose the term (3.3.37) using U-statistics and second order hoeffding projection, particularly the decompositions (2.2.12) – (2.2.14) and (3.2.28) – (3.2.30):

$$
\begin{aligned}
(3.3.37) = &\Bigg(\frac{f(t)}{n^2 h_{1,n}h_{2,n}}\sum_{i=1}^{n}\Big(H_t^g\big((X_i,Y_i),(X_i,Y_i)\big) - \mathbb{E}_{X,Y}H_t^g\big((X_i,Y_i),(X,Y)\big)\Big) \\
&+ \frac{(n-1)f(t)}{nh_{1,n}h_{2,n}}U_n\left(\pi_2(H_t^g(\cdot,\cdot))\right)\Bigg) && (3.3.42) \\
&- \frac{g(t)}{n^2 h_{1,n}h_{2,n}}\sum_{i=1}^{n}(H_t(X_i,X_i) - \mathbb{E}_Z H_t(X_i,Z)) - \frac{(n-1)g(t)}{nh_{1,n}h_{2,n}}U_n\left(\pi_2(H_t(\cdot,\cdot))\right) \\
& && (3.3.43) \\
&+ \frac{n-1}{n^2 h_{1,n}h_{2,n}}\Bigg[f(t)\sum_{i=1}^{n}\Big(\mathbb{E}_{X,Y}H_t^g\big((X,Y),(X_i,Y_i)\big) - \mathbb{E}H_t^g\Big) \\
&- g(t)\sum_{i=1}^{n}(\mathbb{E}_X H_t(X,X_i) - \mathbb{E}H_t)\Bigg]. && (3.3.44)
\end{aligned}
$$

Also, by similar argument as in (3.31) and (3.33) of [10],

$$
\begin{aligned}
\sup_{t\in\mathbb{R}}\Bigg|&(3.3.42) + (3.3.43) - \frac{1}{n^2 h_{1,n}h_{2,n}}\Big[f(t)\sum_{i=1}^{n}\Big(\mathbb{E}_{X,Y}H_t^g\big((X,Y),(X_i,Y_i)\big) - \mathbb{E}H_t^g\Big) \\
&- g(t)\sum_{i=1}^{n}(\mathbb{E}_X H_t(X,X_i) - \mathbb{E}H_t)\Big]\Bigg| = o_{a.s}(h_{2,n}^4). && (3.3.45)
\end{aligned}
$$

The remaining terms in (3.3.37) are *i.i.d.*. By the notation of $P_{n,1}$ in Lemma 3.3.4, note that

$$
P_{n,i} = \frac{f(t)}{h_{1,n}h_{2,n}}\Big(\mathbb{E}_{X,Y}H_t^g\big((X,Y),(X_i,Y_i)\big) - \mathbb{E}H_t^g\Big) - \frac{g(t)}{h_{1,n}h_{2,n}}(\mathbb{E}_X H_t(X,X_i) - \mathbb{E}H_t),
$$

and $T_{n,i} = W_{n,i} + P_{n,i}$. Then, by (3.3.36), (3.3.40), (3.3.41), and (3.3.45) we have

$$
\frac{1}{n}\sum_{i=1}^{n}\hat{W}_{n,i} = \frac{1}{n}\sum_{i=1}^{n}T_{n,i} + o_{a.s}(h_{2,n}^4). \qquad (3.3.46)
$$

Using Theorem 2.4.2 and Lemma 3.3.4, we obtain

$$\sqrt{nh_{2,n}}\left[\frac{1}{n}\sum_{i=1}^{n}\hat{W}_{n,i}-\mathbb{E}(\hat{W}_{n,1})\right]\xrightarrow{D}N\left(0,\sigma_{u1}^2\right). \tag{3.3.47}$$

For $t \in D_{rf}$, $f(t) > 0$. Therefore, by Theorem 1 of [10], $\frac{1}{\hat{f}(t;h_{1,n}h_{2,n})} \xrightarrow{a.s} \frac{1}{f(t)}$. Hence, by (3.3.47) and Theorem 2.4.1, we have

$$\sqrt{nh_{2,n}}[\hat{r}(t;h_{1,n},h_{2,n})-r(t)]\xrightarrow{D}N\left(\frac{\lambda q(t)}{f^2(t)},\sigma_{r1}^2\right).$$

□

4 CONCLUSION

McKay [17] proposed the multidimensional form of the variable bandwidth kernel density estimator (1.2.1) with clipping procedure. In his dissertation, he found the bias form of the ideal estimator and proposed a plug-in true estimator of (1.2.1). For the corresponding true plug-in estimator (1.2.8), Giné & Sang [10] studied the uniform convergence of the estimator in almost sure sense.

In this dissertation, we found the exact form of variance of the ideal estimator proposed by McKay [17]. We worked on the mean squared error of the true variable bandwidth kernel estimator (1.2.8). The convergence rate for the mean square error of the true variable bandwidth density estimator is faster than the classical kernel density estimator. For a fixed t, we developed the central limit theorem for the ideal and true variable bandwidth kernel density estimators in the multidimensional case. Based on the variable bandwidth kernel density estimator studied in this dissertation, we proposed a new estimator for the regression function for one dimensional case. The order of the bias of this new variable bandwidth regression estimator $(O(h_{2,n}^4))$ is better than Nadaraya-Watson estimator $(O((h_{1,n}^2))$. Furthermore, we developed the order of the variance of the ideal and true variable bandwidth kernel regression estimators $(O(nh_{2,n})^{-1})$ which is the same as that of Nadaraya-Watson estimator. Finally, for a fixed point t, we established the central limit theorem for the ideal and true variable kernel regression estimators.

BIBLIOGRAPHY

[1] I. Abramson, *On bandwidth variation in kernel estimates - a square-root law*, Ann. Statist. **10** (1982), 1217–1223.

[2] Herman J. Bierens, *Uniform consistency of kernel estimators of a regression function under generalized conditions*, Journal of the American Statistical Association **79** (1983), no. 383, 699–707.

[3] Patrick Billingsley, *Probability and measure*, Wiley, 1979.

[4] Gérard Collomb, *Estimation non paramétrique de la régression par la méthode du noyau : propriété de convergence asymptotiquememt normale indépendante*, Annales de la Facultédes sciences de l'Université de Clermont. Série Mathématiques des sciences de l'Université de Clermont. Série Mathématiques **65** (1977), 24–46.

[5] ______, *Propriétés de convergence presque complée du prédicteur é noyau*, Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete **66** (1984), 441–460.

[6] U. Einmahl and D.M. Mason, *An empirical process approach to the uniform consistency of kernel-type function estimators*, Journal of Theoretical Probability **13** (2000), 1–37.

[7] ______, *Uniform in bandwidth consistency of kernel-type function estimators*, Ann. Statist. **33** (2005), 1380–1403.

[8] E. Giné and A. Guillou, *Rates of strong uniform consistency for multivariate kernel density estimators*, Ann. I. H. Poincaré-Probab. et Stat. **38** (2002), 907–921.

[9] E. Giné and H. Sang, *Uniform asymptotics for kernel density estimators with variable bandwidths*, J. Nonparametric Statistics **22** (2010), 773–795.

[10] ______, *On the estimation of smooth densities by strict probability densities at optimal rates in sup-norm*, IMS Collections, From Probability to Statistics and Back: High-Dimensional Models and Processes **9** (2013), 128–149.

[11] P. Hall, *On the bias of variable bandwidth kernel estimators*, Biometrika **77** (1990), 529–535.

[12] P. Hall, T. Hu, and J. S. Marron, *Improved variable window kernel estimates of probability densities*, Ann. Statist. **23** (1995), 1–10.

[13] P. Hall and J. S. Marron, *Variable window width kernel estimates of probability densities*, Probab. Th. Rel. Fields **80** (1988), 37–49.

[14] Robert V. Hogg, Joseph W. McKean, and Allen Thornton Craig, *Introduction to mathematical statistics*, Pearson, 2013.

[15] M. C. Jones, I. J. McKay, and T.C. Hu, *Variable location and scale kernel density estimation*, Ann. Inst. Statist. Math **46** (1994), 521–535.

[16] I. J. McKay, *A note on bias reduction in variable kernel density estimates*, Canad. J. Statist. **21** (1993), 367–375.

[17] ———, *Variable kernel methods in density estimation*, Ph.D. thesis, Queen's University, 1993.

[18] H. Müller and U Stadtmüller, *Variable bandwidth kernel estimators of regression curves*, Ann. Statist. **15** (1987), 182–201.

[19] E. A. Nadaraya, *On estimating regression*, Theory of Probability and Its Applications **9** (1964), no. 1, 141–142.

[20] Kazuo Noda, *Estimation of a regression function by the parzen kernel-type density estimators*, Annals of the Institute of Statistical Mathematics **28** (1976), no. 1, 221–234.

[21] S. Yu. Novak, *A generalized kernel density estimator*, Theory Probab. Appl. **44** (2000), 570–583.

[22] Emanuel Parzen, *On estimation of a probability density function and mode*, Annals of Mathematical Statistics **33** (1962), no. 3, 1065–1076.

[23] M. Samiuddin and G. M. El-Sayyad, *On nonparametric kernel density estimates*, Biometrika **77** (1990), no. 4, 865–874.

[24] Michael G. Schimek, *Smoothing and regression : Approaches, computation, and application.*, John Wiley & Sons, 2000.

[25] B. W. Silverman, *Density estimation for statistics and data analysis*, Chapman and Hall, 1986.

[26] G. R. Terrell and D. Scott, *Variable kernel density estimation*, Ann. Statist **20** (1992), 1236–1265.

[27] Larry Wasserman, *All of nonparametric statistics*, Springer, 2006.

[28] G. S. Watson, *Smooth regression analysis*, Sankhya. Series A **26** (1964), no. 4, 359–372.

VITA

Janet Nakarmi was born in Kathmandu, Nepal to Narayan Lal and Purna Laxmi Nakarmi. She received a Bachelor's of Science degree in 2010 from Randolph College and a Master's of Science degree in 2013 from The University of Mississippi. She is currently a Ph.D. candidate in Mathematics at The University of Mississippi.